

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 04:08:02 ; Search time 64.5137 Seconds  
(without alignments)  
1990.064 Million cell updates/sec

Title: US-10-631-958-2

Perfect score: 1717

Sequence: 1 PKHLVFNPGGKQGRKRI.....KCSRNFRLRIHTNQDQ 326

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1774312 seqs, 393823214 residues

Total number of hits satisfying chosen parameters: 1774312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1717	100.0	326	9 US-09-784-810A-11	Sequence 11, Appl
2	1717	100.0	326	10 US-09-969-896-2	Sequence 2, Appli
3	1717	100.0	326	16 US-10-631-958-2	Sequence 2, Appli
4	1717	100.0	326	18 US-10-876-281-11	Sequence 11, Appl
5	1717	100.0	471	9 US-09-784-810A-6	Sequence 6, Appli
6	1717	100.0	471	18 US-10-876-281-6	Sequence 6, Appli
7	1640.5	95.5	537	10 US-09-969-896-10	Sequence 10, Appl
8	1640.5	95.5	537	15 US-10-262-511-40	Sequence 40, Appl
9	1640.5	95.5	537	16 US-10-631-958-10	Sequence 10, Appl
10	1640.5	95.5	537	16 US-10-618-941-121	Sequence 121, App
11	1640.5	95.5	562	10 US-09-969-896-11	Sequence 11, Appl

12	1640.5	95.5	562	16 US-10-631-958-11	Sequence 11, Appl
13	1632.5	95.1	537	14 US-10-325-597A-2	Sequence 2, Appli
14	529	30.8	588	16 US-10-425-115-361076	Sequence 361076,
15	476.5	27.8	182	9 US-09-784-810A-29	Sequence 29, Appl
16	476.5	27.8	182	18 US-10-876-281-29	Sequence 29, Appl
17	452	26.3	575	16 US-10-425-115-270874	Sequence 270874,
18	335	19.5	484	16 US-10-437-963-133133	Sequence 133133,
19	334	19.5	245	15 US-10-425-114-53957	Sequence 53957, A
20	333	19.4	487	15 US-10-424-599-190238	Sequence 190238,
21	327	19.0	490	16 US-10-425-115-249553	Sequence 249553,
22	321.5	18.7	613	16 US-10-437-963-195284	Sequence 195284,
23	318	18.5	760	16 US-10-425-115-213665	Sequence 213665,
24	311.5	18.1	381	16 US-10-425-115-249555	Sequence 249555,
25	306	17.8	520	15 US-10-424-599-202261	Sequence 202261,
26	305.5	17.8	470	16 US-10-425-115-282512	Sequence 282512,
27	305.5	17.8	521	15 US-10-425-114-68423	Sequence 68423, A
28	300	17.5	446	16 US-10-437-963-197631	Sequence 197631,
29	293	17.1	677	16 US-10-425-115-264087	Sequence 264087,
30	289.5	16.9	373	16 US-10-767-701-44927	Sequence 44927, A
31	288	16.8	764	16 US-10-437-963-135389	Sequence 135389,
32	287.5	16.7	757	16 US-10-437-963-180379	Sequence 180379,
33	282	16.4	490	14 US-10-053-510-19	Sequence 19, Appl
34	282	16.4	490	15 US-10-348-052-19	Sequence 19, Appl
35	282	16.4	490	15 US-10-622-011-19	Sequence 19, Appl
36	282	16.4	641	15 US-10-348-052-28	Sequence 28, Appl
37	282	16.4	641	16 US-10-622-011-28	Sequence 28, Appl
38	273	15.9	144	9 US-09-784-810A-8	Sequence 8, Appli
39	273	15.9	144	18 US-10-876-281-8	Sequence 20, Appl
40	270.5	15.8	524	14 US-10-053-510-20	Sequence 20, Appl
41	270.5	15.8	524	15 US-10-348-052-20	Sequence 20, Appl
42	270.5	15.8	524	16 US-10-622-011-20	Sequence 20, Appl
43	270.5	15.8	907	15 US-10-348-052-29	Sequence 29, Appl
44	270.5	15.8	907	16 US-10-622-011-29	Sequence 29, Appl
45	269.5	15.7	552	16 US-10-437-963-161184	Sequence 161184,

ALIGNMENTS

RESULT 1

US-09-784-810A-11  
; Sequence 11, Application US/09784810A  
; Patent No. US20020082203A1

; GENERAL INFORMATION:

; APPLICANT: RASTELLI, LUCA

; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING

; FILE REFERENCE: 10716-08

; CURRENT APPLICATION NUMBER: US/09/784,810A

; CURRENT FILING DATE: 2001-02-14

; PRIOR APPLICATION NUMBER: 60/182,360

; PRIOR FILING DATE: 2000-02-14

; PRIOR APPLICATION NUMBER: 60/191,261

; PRIOR FILING DATE: 2000-03-22

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 11

; LENGTH: 326

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-784-810A-11

Query Match 100.0%; Score 1717; DB 9; Length 326;

Best Local Similarity 100.0%; Pred. No. 2.8e-172;

Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKHLVFNPGGKQGRKRIYKRVAPLFTLASITTDIIGNKFYNYVEVITEHANQAKE 60

Db 1 PKHLVFNPGGKQGRKRIYKRVAPLFTLASITTDIIGNKFYNYVEVITEHANQAKE 60

QY 61 TLYEINIDKYGIVCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPA 120

Db 61 TLYEINIDKYGIVCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPA 120

Qy 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSHTLLRYSVSLGCVGFVGDIIK 180  
Db 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSHTLLRYSVSLGCVGFVGDIIK 180  
Qy 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRCAGCFVCRQSK 240  
Db 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRCAGCFVCRQSK 240  
Qy 241 QOLEEQQKALYGLEAAEDVEEQVVCCKFLAINATNMNSCACRSPRGLSPAHLGDGSS 300  
Db 241 QOLEEQQKALYGLEAAEDVEEQVVCCKFLAINATNMNSCACRSPRGLSPAHLGDGSS 300  
Qy 301 DLILIRKSRNFNLFRLIRHTNQDQ 326  
Db 301 DLILIRKSRNFNLFRLIRHTNQDQ 326

## RESULT 2

US-09-969-896-2  
; Sequence 2, Application US/09969896  
; Publication No. US20030125533A1

; GENERAL INFORMATION:  
; APPLICANT: Kossida, Sophia  
; TITLE OF INVENTION: Regulation of human Sphingosine  
; FILE REFERENCE: 004974.00594  
; CURRENT APPLICATION NUMBER: US/09/969,896  
; PRIOR FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: US 60/238,005  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: US 60/314,113  
; PRIOR FILING DATE: 2001-08-23  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2  
; LENGTH: 326  
; TYPE: PRT

ORGANISM: Homo sapiens  
US-09-969-896-2

Query Match 100.0%; Score 1717; DB 10; Length 326;  
Best Local Similarity 100.0%; Pred. No. 2.8e-172;  
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVYVEVITEHANQAKE 60  
Db 1 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVYVEVITEHANQAKE 60  
Qy 61 TLYEINIDKYDGI VCVGDDGMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPA 120  
Db 61 TLYEINIDKYDGI VCVGDDGMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPA 120  
Qy 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSHTLLRYSVSLGCVGFVGDIIK 180  
Db 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSHTLLRYSVSLGCVGFVGDIIK 180  
Qy 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRCAGCFVCRQSK 240  
Db 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRCAGCFVCRQSK 240  
Qy 241 QOLEEQQKALYGLEAAEDVEEQVVCCKFLAINATNMNSCACRSPRGLSPAHLGDGSS 300  
Db 241 QOLEEQQKALYGLEAAEDVEEQVVCCKFLAINATNMNSCACRSPRGLSPAHLGDGSS 300  
Qy 301 DLILIRKSRNFNLFRLIRHTNQDQ 326  
Db 301 DLILIRKSRNFNLFRLIRHTNQDQ 326

## RESULT 3

US-10-631-958-2  
; Sequence 2, Application US/10631958

; Publication No. US20040192580A1  
; GENERAL INFORMATION:  
; APPLICANT: Kossida, Sophia  
; TITLE OF INVENTION: Regulation of human Sphingosine  
; FILE REFERENCE: 004974.00594  
; CURRENT APPLICATION NUMBER: US/10/631,958  
; CURRENT FILING DATE: 2003-08-01  
; PRIOR APPLICATION NUMBER: US/09/969,896  
; PRIOR FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: US 60/238,005  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: US 60/314,113  
; PRIOR FILING DATE: 2001-08-23  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 326  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-631-958-2

Query Match 100.0%; Score 1717; DB 16; Length 326;  
Best Local Similarity 100.0%; Pred. No. 2.8e-172;  
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVYVEVITEHANQAKE 60  
Db 1 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVYVEVITEHANQAKE 60  
Qy 61 TLYEINIDKYDGI VCVGDDGMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPA 120  
Db 61 TLYEINIDKYDGI VCVGDDGMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPA 120  
Qy 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSHTLLRYSVSLGCVGFVGDIIK 180  
Db 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSVVHNSHTLLRYSVSLGCVGFVGDIIK 180  
Qy 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRCAGCFVCRQSK 240  
Db 181 DSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRCAGCFVCRQSK 240  
Qy 241 QOLEEQQKALYGLEAAEDVEEQVVCCKFLAINATNMNSCACRSPRGLSPAHLGDGSS 300  
Db 241 QOLEEQQKALYGLEAAEDVEEQVVCCKFLAINATNMNSCACRSPRGLSPAHLGDGSS 300  
Qy 301 DLILIRKSRNFNLFRLIRHTNQDQ 326  
Db 301 DLILIRKSRNFNLFRLIRHTNQDQ 326

## RESULT 4

US-10-876-281-11  
; Sequence 11, Application US/10876281  
; Publication No. US20050123942A1  
; GENERAL INFORMATION:  
; APPLICANT: RASTELLI, LUCA  
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING  
; FILE REFERENCE: 10716-08  
; CURRENT APPLICATION NUMBER: US/10/876,281  
; CURRENT FILING DATE: 2004-06-24  
; PRIOR APPLICATION NUMBER: US/09/784,810  
; PRIOR FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: 60/182,360  
; PRIOR FILING DATE: 2000-02-14  
; PRIOR APPLICATION NUMBER: 60/191,261  
; PRIOR FILING DATE: 2000-03-22  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 326  
; TYPE: PRT



```
; ORGANISM: Homo sapiens
US-10-876-281-11

Query Match      100.0%; Score 1717; DB 18; Length 326;
Best Local Similarity 100.0%; Pred. No. 2.8e-172;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKHLVFINPFGGKGQKRIYERKVAFLFTLASITTTDIIIGNKFYVYVEVITEHANOAKE 60
DB 1 PKHLVFINPFGGKGQKRIYERKVAFLFTLASITTTDIIIGNKFYVYVEVITEHANOAKE 60

QY 61 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPA 120
DB 61 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPA 120

QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSVVHNSLTLLRYSVSLGFGFYGDIK 180
DB 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSVVHNSLTLLRYSVSLGFGFYGDIK 180

QY 181 DSEKKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRQSK 240
DB 181 DSEKKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRQSK 240

QY 241 QOLEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300
DB 241 QOLEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300

QY 301 DLILIRKCSRFNFLRLIRHTNQDDQ 326
DB 301 DLILIRKCSRFNFLRLIRHTNQDDQ 326

QY 301 DLILIRKCSRFNFLRLIRHTNQDDQ 326
DB 301 DLILIRKCSRFNFLRLIRHTNQDDQ 326

RESULT 5
US-09-784-810A-6
; Sequence 6, Application US/09784810A
; Patent No. US20020082203A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/09/784,810A
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-784-810A-6

Query Match      100.0%; Score 1717; DB 9; Length 471;
Best Local Similarity 100.0%; Pred. No. 4.7e-172;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKHLVFINPFGGKGQKRIYERKVAFLFTLASITTTDIIIGNKFYVYVEVITEHANOAKE 60
DB 53 PKHLVFINPFGGKGQKRIYERKVAFLFTLASITTTDIIIGNKFYVYVEVITEHANOAKE 112

QY 61 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPA 120
DB 113 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPA 172

QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSVVHNSLTLLRYSVSLGFGFYGDIK 180
DB 173 GSTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSVVHNSLTLLRYSVSLGFGFYGDIK 232

QY 181 DSEKKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRQSK 240
DB 181 DSEKKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRQSK 240

Query Match      100.0%; Score 1717; DB 9; Length 471;
Best Local Similarity 100.0%; Pred. No. 4.7e-172;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKHLVFINPFGGKGQKRIYERKVAFLFTLASITTTDIIIGNKFYVYVEVITEHANOAKE 60
DB 53 PKHLVFINPFGGKGQKRIYERKVAFLFTLASITTTDIIIGNKFYVYVEVITEHANOAKE 112

QY 61 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPA 120
DB 113 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPA 172

QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSVVHNSLTLLRYSVSLGFGFYGDIK 180
DB 173 GSTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSVVHNSLTLLRYSVSLGFGFYGDIK 232

QY 181 DSEKKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRQSK 240
DB 181 DSEKKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRQSK 240

; ORGANISM: Homo sapiens
US-09-784-810A-6

Query Match      100.0%; Score 1717; DB 18; Length 471;
Best Local Similarity 100.0%; Pred. No. 4.7e-172;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKHLVFINPFGGKGQKRIYERKVAFLFTLASITTTDIIIGNKFYVYVEVITEHANOAKE 60
DB 53 PKHLVFINPFGGKGQKRIYERKVAFLFTLASITTTDIIIGNKFYVYVEVITEHANOAKE 112

QY 61 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPA 120
DB 113 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPA 172

QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSVVHNSLTLLRYSVSLGFGFYGDIK 180
DB 173 GSTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSVVHNSLTLLRYSVSLGFGFYGDIK 232

QY 181 DSEKKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRQSK 240
DB 181 DSEKKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRQSK 240

; ORGANISM: Homo sapiens
US-10-876-281-6
; Sequence 6, Application US/10876281
; Publication No. US20050123942A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/10/876,281
; PRIOR FILING DATE: 2004-06-24
; PRIOR APPLICATION NUMBER: US/09/784,810
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-876-281-6

Query Match      100.0%; Score 1717; DB 18; Length 471;
Best Local Similarity 100.0%; Pred. No. 4.7e-172;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKHLVFINPFGGKGQKRIYERKVAFLFTLASITTTDIIIGNKFYVYVEVITEHANOAKE 60
DB 53 PKHLVFINPFGGKGQKRIYERKVAFLFTLASITTTDIIIGNKFYVYVEVITEHANOAKE 112

QY 61 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPA 120
DB 113 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPA 172

QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSVVHNSLTLLRYSVSLGFGFYGDIK 180
DB 173 GSTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSVVHNSLTLLRYSVSLGFGFYGDIK 232

QY 181 DSEKKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRQSK 240
DB 181 DSEKKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRQSK 240

QY 241 QOLEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300
DB 241 QOLEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300

QY 293 QOLEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 352
DB 293 QOLEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 352

QY 301 DLILIRKCSRFNFLRLIRHTNQDDQ 326
DB 301 DLILIRKCSRFNFLRLIRHTNQDDQ 326

QY 301 DLILIRKCSRFNFLRLIRHTNQDDQ 326
DB 301 DLILIRKCSRFNFLRLIRHTNQDDQ 326

RESULT 7
US-09-896-896-10
; Sequence 10, Application US/09969896
; Publication No. US2003012533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; Kinase-Like Protein
```



; TITLE OF INVENTION: Regulation of human Sphingosine  
 ; FILE REFERENCE: 004974.00594  
 ; CURRENT APPLICATION NUMBER: US/10/631,958  
 ; PRIOR FILING DATE: 2003-08-01  
 ; PRIOR APPLICATION NUMBER: US/09/969,896  
 ; PRIOR FILING DATE: 2001-10-04  
 ; PRIOR APPLICATION NUMBER: US 60/238,005  
 ; PRIOR FILING DATE: 2000-10-06  
 ; PRIOR APPLICATION NUMBER: US 60/314,113  
 ; PRIOR FILING DATE: 2001-08-23  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 10  
 ; LENGTH: 537  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-631-958-10

Query Match 95.5%; Score 1640.5; DB 16; Length 537;  
 Best Local Similarity 96.3%; Pred. No. 6.9e-164;  
 Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;  
 QY 1 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVYVEVITEHANOAKE 60  
 DB 130 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 178  
 QY 61 TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPA 120  
 DB 179 TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPA 238  
 QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHNSLTLRYSVSLLGYGYGDIK 180  
 DB 239 GSTDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHNSLTLRYSVSLLGYGYGDIK 298  
 QY 181 DSEKKRWGLARYDFSGKTLFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRSK 240  
 DB 299 DSEKKRWGLARYDFSGKTLFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRSK 358  
 QY 241 QOLEEQQKALYGLEAAEDVEEWQVCGKFLAINATNNSCACRRSPRGLSPAHLGDGSS 300  
 DB 359 QOLEEQQKALYGLEAAEDVEEWQVCGKFLAINATNNSCACRRSPRGLSPAHLGDGSS 418  
 QY 301 DLILIRKSRFNLRLIRHTNQDQ 326  
 DB 419 DLILIRKSRFNLRLIRHTNQDQ 444

RESULT 10  
 US-10-618-941-121  
 ; Sequence 121, Application US/10618941  
 ; Publication No. US20040197792A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WHYTE, DAVID  
 ; APPLICANT: MANNING, GERARD  
 ; APPLICANT: CAENEPEEL, SEAN  
 ; TITLE OF INVENTION: NOVEL KINASES  
 ; FILE REFERENCE: 034536-0321  
 ; CURRENT APPLICATION NUMBER: US/10/618,941  
 ; CURRENT FILING DATE: 2003-07-15  
 ; PRIOR APPLICATION NUMBER: 60/395,632  
 ; PRIOR FILING DATE: 2002-07-15  
 ; NUMBER OF SEQ ID NOS: 143  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 121  
 ; LENGTH: 537  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-618-941-121

Query Match 95.5%; Score 1640.5; DB 16; Length 537;  
 Best Local Similarity 96.3%; Pred. No. 6.9e-164;  
 Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

QY 1 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVYVEVITEHANOAKE 60  
 DB 130 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 178  
 QY 61 TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPA 120  
 DB 179 TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPA 238  
 QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHNSLTLRYSVSLLGYGYGDIK 180  
 DB 239 GSTDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHNSLTLRYSVSLLGYGYGDIK 298  
 QY 181 DSEKKRWGLARYDFSGKTLFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRSK 240  
 DB 299 DSEKKRWGLARYDFSGKTLFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRSK 358  
 QY 241 QOLEEQQKALYGLEAAEDVEEWQVCGKFLAINATNNSCACRRSPRGLSPAHLGDGSS 300  
 DB 359 QOLEEQQKALYGLEAAEDVEEWQVCGKFLAINATNNSCACRRSPRGLSPAHLGDGSS 418  
 QY 301 DLILIRKSRFNLRLIRHTNQDQ 326  
 DB 419 DLILIRKSRFNLRLIRHTNQDQ 444

RESULT 11  
 US-09-969-896-11  
 ; Sequence 11, Application US/09969896  
 ; Publication No. US20030125533A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kossida, Sophia  
 ; TITLE OF INVENTION: Regulation of human Sphingosine  
 ; FILE REFERENCE: 004974.00594  
 ; CURRENT APPLICATION NUMBER: US/09/969,896  
 ; CURRENT FILING DATE: 2001-10-04  
 ; PRIOR APPLICATION NUMBER: US 60/238,005  
 ; PRIOR FILING DATE: 2000-10-06  
 ; PRIOR APPLICATION NUMBER: US 60/314,113  
 ; PRIOR FILING DATE: 2001-08-23  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 11  
 ; LENGTH: 562  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-969-896-11

Query Match 95.5%; Score 1640.5; DB 10; Length 562;  
 Best Local Similarity 96.3%; Pred. No. 7.4e-164;  
 Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;  
 QY 1 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVYVEVITEHANOAKE 60  
 DB 155 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 203  
 QY 61 TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPA 120  
 DB 204 TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPA 263  
 QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHNSLTLRYSVSLLGYGYGDIK 180  
 DB 264 GSTDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHNSLTLRYSVSLLGYGYGDIK 323  
 QY 181 DSEKKRWGLARYDFSGKTLFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRSK 240  
 DB 324 DSEKKRWGLARYDFSGKTLFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRACGFCVCRSK 383  
 QY 241 QOLEEQQKALYGLEAAEDVEEWQVCGKFLAINATNNSCACRRSPRGLSPAHLGDGSS 300  
 DB 384 QOLEEQQKALYGLEAAEDVEEWQVCGKFLAINATNNSCACRRSPRGLSPAHLGDGSS 443

Qy 301 DLILIRKSRFNFRLRLIRHTNQDQ 326  
Db 444 DLILIRKSRFNFRLRLIRHTNQDQ 469

## RESULT 12

US-10-631-958-11  
; Sequence 11, Application US/10631958  
; Publication No. US20040192580A1  
; GENERAL INFORMATION:  
; APPLICANT: Kossida, Sophia  
; TITLE OF INVENTION: Regulation of human Sphingosine  
; FILE REFERENCE: 004974.00594  
; CURRENT APPLICATION NUMBER: US/10/631,958  
; CURRENT FILING DATE: 2003-08-01  
; PRIOR APPLICATION NUMBER: US/09/969,896  
; PRIOR FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: US 60/238,005  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: US 60/314,113  
; PRIOR FILING DATE: 2001-08-23  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 562  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-631-958-11

Query Match 95.5%; Score 1640.5; DB 16; Length 562;  
Best Local Similarity 96.3%; Pred. No. 7.4e-164;  
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

Qy 1 PKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVYVEVITEHANQAKE 60  
Db 155 PKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDII-----VTEHANQAKE 203  
Qy 61 TLYEINIDKYDGI VCVGDMGFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPA 120  
Db 204 TLYEINIDKYDGI VCVGDMGFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPA 263  
Qy 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSSVHHNSTLLRYSVSLGCGFYGDIK 180  
Db 264 GSTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSSVHHNSTLLRYSVSLGCGFYGDIK 323  
Qy 181 DSEKRWGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCRQSK 240  
Db 324 DSEKRWGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCRQSK 383  
Qy 241 QOLEEOKKALYGLEAAEDVEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300  
Db 384 QOLEEOKKALYGLEAAEDVEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 443  
Qy 301 DLILIRKSRFNFRLRLIRHTNQDQ 326  
Db 444 DLILIRKSRFNFRLRLIRHTNQDQ 469

## RESULT 13

US-10-315-597A-2  
; Sequence 2, Application US/10315597A  
; Publication No. US20030162206A1  
; GENERAL INFORMATION:  
; APPLICANT: Sugliura, Masako  
; APPLICANT: Kono, Keita  
; APPLICANT: Kohama, Takafumi  
; TITLE OF INVENTION: Ceramide Kinase and DNA Encoding It  
; FILE REFERENCE: 02658CIP/HG  
; CURRENT APPLICATION NUMBER: US/10/315,597A  
; CURRENT FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: JP 2000-178039  
; PRIOR FILING DATE: 2000-06-14

; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 537  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-315-597A-2

Query Match 95.1%; Score 1632.5; DB 14; Length 537;  
Best Local Similarity 95.7%; Pred. No. 4.8e-163;  
Matches 312; Conservative 2; Mismatches 1; Indels 11; Gaps 1;

Qy 1 PKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVYVEVITEHANQAKE 60  
Db 130 PKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDII-----VTEHANQAKE 178  
Qy 61 TLYEINIDKYDGI VCVGDMGFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPA 120  
Db 179 TLYEINIDKYDGI VCVGDMGFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPA 238  
Qy 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSSVHHNSTLLRYSVSLGCGFYGDIK 180  
Db 239 GPTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSSVHHNSTLLRYSVSLGCGFYGDIK 298  
Qy 181 DSEKRWGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCRQSK 240  
Db 299 DSEKRWGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCRQSK 358  
Qy 241 QOLEEOKKALYGLEAAEDVEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300  
Db 359 QOLEEOKKALYGLEAAEDVEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 418  
Qy 301 DLILIRKSRFNFRLRLIRHTNQDQ 326  
Db 419 DLILIRKSRFNFRLRLIRHTNQDQ 444

## RESULT 14

US-10-425-115-361076  
; Sequence 361076, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 361076  
; LENGTH: 588  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(588)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_92481C.1.pap  
US-10-425-115-361076

Query Match 30.8%; Score 529; DB 16; Length 588;  
Best Local Similarity 32.0%; Pred. No. 2.2e-46;  
Matches 131; Conservative 57; Mismatches 107; Indels 114; Gaps 11;

Qy 1 PKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVYVEVITEHANQAKE 60  
Db 186 PKNLMVFHPLCGKRGVSNWE-TYVPLFAKAKTKVI-----VTERAGHAYD 233  
Qy 61 T---LYEINIDKYDGI VCVGDMGFSEVLHGLIGRTOR----- 95

```
Db      234 TLSSLSLDELKAFDGVAVGGDLGFLNEILGILSSRHKVSYPPTPEGFGYVGSSDNCREQ 293
QY      96  -----SAGVDQ-----NHPRAVLPSS----- 112
Db      294 TRNMWNFSKPTPDGSGNAVFLGNPSKDDQDEPLLTSTSPSELEIPSNSTKEPCIGQDNPV 353
QY     113 -----LRGIIIPAGSTDCVCTVGTSDAETSALHIVVGDSLAMDVSVVHNST----- 161
Db     354 SPFNDFRUGIIPSGSTDAIVLSTTGERDPVTSAIIILGRKVSLEDAQVVRWKTSPAAR 413
QY     162 ---LLRYSVSLGCGFYGBDIIKDSEKKRWGLARYDFGLKTFLSHHCVGTVSFLPAQH 218
Db     414 VLPTRYAASPAAGYGFYGEVIRESEKYRWGPARYDFSGTMVFLKHSYGAKVAFDTPY 473
QY     219 T---VGSPRD-----RKP-----CRAGCFVCRSQKQLEBEEQKALYGLEAAED 259
Db     474 THSLTASAQDDITGAQLQSRWKKPKRIICRTNCFVCKEASTSGQNPDDVADNSRTICEN 533
QY     260 VVEWQVVCCKFLAINATNMSCACRSRGLSPAHLGDSDDLIIIRKC 308
Db     534 -QKWMSEGHFLSVGAATVSCRNERAPDGLVADAHLSGDLHLLLRDC 581
```

## RESULT 15

```
US-09-784-810A-29
; Sequence 29, Application US/09784810A
; Patent No. US20020082203A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: SAME
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/09/784,810A
; CURRENT FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: 80432911
US-09-784-810A-29
```

```
Query Match      27.8%; Score 476.5; DB 9; Length 182;
Best Local Similarity 88.6%; Pred. No. 1.4e-41;
Matches 93; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

QY      1 PKHLLVFNPFGGKGKGRYERKVAPLFTLASITTDIIGNKFYVYVVEVITEHANQAKE 60
Db     89 PKHLLVFNPFGGKGKGRYERKVAPLFTLASITTDII-----VTEHANQAKE 137
QY     61 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPR 105
Db    138 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPR 182
```

Search completed: September 3, 2005, 04:32:26  
Job time : 66.5137 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: September 6, 2005, 02:19:54 ; Search time 503.298 Seconds  
(without alignments)  
3834.378 Million cell updates/sec

Title: US-10-631-958-2  
Perfect score: 1717  
Sequence: 1 PKHLVFINPFGKCGQKRI.....KCSRFNFLRLIRHTNQDQ 326

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues  
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cn2.1/USPTO.spool/US10631958/runat.02092005.165814.4497/app.query.fasta\_1.1941  
-DB=N\_Geneseq.16Dec04 -QFMT=fastap -SUFIX=ring -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LISTS=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10631958 @CGN 1.1436 @runat.02092005.165814.4497 -NCFU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq.16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1717	100.0	979	6	ABL40822 Human sph
2	1717	100.0	1840	5	Aad14426 Human sph
3	1640.5	95.5	1740	8	Ada05679 Human NOV
4	1640.5	95.5	3975	8	Acc70838 Human sph
5	1640.5	95.5	4231	3	Aaa50510 Human sph

6	1640.5	95.5	4413	6	ABL40828	Abi40828 Human sph
7	1640.5	95.5	4429	12	ADJ96598	Adj96598 Human lip
8	1640.5	95.5	4445	13	ADP55247	Adp55247 Human PRO
9	1633.5	95.1	4432	8	Abx70921	Abx70921 Novel hum
10	1632.5	95.1	4463	6	ABA96945	Abx96945 Human cer
11	1595.5	92.9	1740	12	ADN62844	Adn62844 Human NOV
12	1414.5	82.4	4702	13	ADS10370	Adsl10370 Human the
13	1400.5	81.6	2241	5	AAS77728	Aas77728 DNA encod
14	1323.5	77.1	2186	5	AAS77730	Aas77730 DNA encod
15	1055	61.4	753	3	AACT6031	Aac76031 Human ORF
16	986	57.4	1570	5	AAS77731	Aas77731 DNA encod
17	881	47.2	474	6	ABL40823	Abi40823 Human sph
18	679	39.5	426	5	AAS77727	Aas77727 DNA encod
19	626	36.5	411	5	AAS77729	Aas77729 DNA encod
20	613	35.7	2173	4	ABL25705	Abi25705 Drosophil
21	613	35.7	10337	4	ABL25704	Abi25704 Drosophil
22	508.5	29.6	329	6	ABL40824	Abi40824 Human sph
23	476.5	27.8	547	3	AACT6592	Aac76592 Human ORF
24	442	25.7	1774	12	ADO88890	Ado88890 Novel hum
25	414.5	24.1	564	8	ABT23453	Abt23453 Immune-re
26	389	22.7	15185	4	AAK65589	Aak65589 Human inm
27	386	22.5	15181	4	AAK65588	Aak65588 Human inm
28	302	17.6	167	6	ABL40825	Abi40825 Human sph
29	300.5	17.5	817	6	ABQ99499	Abq99499 Human cod
30	300.5	17.5	817	13	ADS11669	Adsl1669 Human the
31	282	16.4	2020	4	ABL03297	Abi03297 Drosophil
32	282	16.4	2629	9	ACF35859	Acf35859 D. melano
33	282	16.4	4020	4	ABL03296	Abi03296 Drosophil
34	273	15.9	522	5	AAD14427	Aad14427 Partial r
35	270.5	15.8	2609	9	ACF35860	Acf35860 D. melano
36	270.5	15.8	2830	4	ABL02327	Abi02327 Drosophil
37	269	15.7	572	13	ACNS7851	Acns7851 Cotton gy
38	260	15.1	153	6	ABL40826	Abi40826 Human sph
39	243	14.2	1857	6	ABK90199	Abk90199 cDNA enco
40	243	14.2	2380	5	AAS14817	Aas14817 Human cDN
41	243	14.2	2380	6	ABL59533	Abi59533 Human sph
42	243	14.2	2380	10	ADE85298	Ade85298 Farnesyl
43	243	14.2	2380	10	ADE38416	Ade38416 Human can
44	243	14.2	2380	12	ADQ15181	Adq15181 Human can
45	243	14.2	2422	5	AAS85331	Aas85331 DNA encod

ALIGNMENTS

RESULT 1  
ABL40822  
ID ABL40822 standard; cDNA; 979 BP.  
XX  
AC ABL40822;  
XX  
DT 03-JUL-2002 (first entry)  
XX  
DE Human sphingosine kinase-like protein encoding cDNA.  
KW Human sphingosine kinase-like protein; intracellular signalling; gene;  
KW cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;  
KW autoimmune disease; rheumatoid arthritis; Parkinson's disease; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
CDS 2..789  
FT /\*tag= a  
FT /product= "sphingosine kinase-like protein"  
FT /note= "start and stop codons are not indicated"  
XX  
PN WO200228906-A2.  
XX  
PD 11-APR-2002.  
XX  
PF 05-OCT-2001; 2001WO-EP011516.  
XX  
PR 06-OCT-2000; 2000US-0238005P.











XX WO200052173-A2.  
 XX  
 XX  
 XX PD 08-SEP-2000.  
 XX  
 XX PF 02-MAR-2000; 2000WO-CA000223.  
 XX  
 XX PR 02-MAR-1999; 99US-0122516P.  
 XX  
 XX PA (ALLX ) NPS ALLELIX CORP.  
 XX  
 XX PI Munroe D, Gupta A, Falzone GR;  
 XX P-PSDB; AAY96059.  
 XX  
 XX DR WPI; 2000-572185/53.  
 XX  
 XX PT New human sphingosine kinase A, B and C polynucleotides and polypeptides  
 XX useful in e.g. chromosome and gene mapping, and detecting inflammation or  
 XX disease associated with abnormal levels of sphingosine kinase expression.  
 XX  
 XX PS Disclosure; Fig 7; 81pp; English.  
 XX  
 XX CC The present sequence is that of an isolated polynucleotide encoding human  
 XX sphingosine kinase C (SKC, see AAY96059), an enzyme that phosphorylates  
 XX sphingosine to form sphingosine 1-phosphate. The polynucleotide was  
 XX isolated from an HeLa cDNA library by PCR amplification. The invention  
 XX provides polynucleotides (see AAY50508-10) and polypeptides (see AAY96057  
 XX -59) for the human sphingosine kinase (SK) homologues SKA, SKB and SKC.  
 XX The polynucleotides may be used as hybridization probes, in the  
 XX construction of PCR primers for chromosome and gene mapping, in the  
 XX recombinant production of SKA, SKB and SKC, and in the generation of  
 XX antisense DNA or RNA. They can be used to detect inflammation or disease  
 XX associated with abnormal levels of SK expression, or to detect  
 XX differences in gene sequence between normal and carrier or affected  
 XX individuals. Host cells expressing SK can be used in drug screening.  
 XX Human SK specific antibodies, inhibitors, ligands or their analogues are  
 XX useful as bioactive agents to treat inflammation or disease including  
 XX viral, bacterial or fungal infections, allergic responses, mechanical  
 XX injury associated with trauma, hereditary diseases, lymphoma or  
 XX carcinoma, and other conditions with activate the genes of kidney, lung,  
 XX heart, lymphoid or tissues of the nervous system  
 XX  
 XX SQ Sequence 4231 BP; 1022 A; 1021 C; 1086 G; 1067 T; 0 U; 35 Other;

Alignment Scores:  
 Pred. No.: 4,63e-170 Length: 4231  
 Score: 1640.50 Matches: 314  
 Percent Similarity: 96.63% Conservative: 1  
 Best Local Similarity: 96.32% Mismatches: 0  
 Query Match: 95.54% Indels: 11  
 DB: 3 Gaps: 1

US-10-631-958-2 (1-326) x AAY50510 (1-4231)

Qy 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyGlyGlnGlyLysArgile 20  
 Db 227 CCARAGCATTACTGCTATTATCAACCGTTTGGAGAAAGGACAGGCGGATA 286  
 Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40  
 Db 287 TATGAAAGAAAGTGGCCACCACTGTTACCTTAGCCTCCATCACCACCTGACATCATC--- 343  
 Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60  
 Db 344 -----GTTACTGAACATGCTTAATCAGGCCAAGGAG 373  
 Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyLysAspGly 80  
 Db 374 ACTCTGTATGAGATTAAATACACAAATACGCGCATCGTCTGTGCGGGAGATGTT 433  
 Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100  
 Db 434 ATGTTTCAGCGAGTCTGCACGGTCTGATTGGGAGGACGAGGAGCGCGGGGTCGAC 493

Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120  
 Db 494 CAGAACACACCCCGGGCTGTGCTGCCAGTAGACCTCGGATTTGGATCATTTCCGCA 553  
 Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140  
 Db 554 GGGTCAACGGGACTGCGTGTGTACTCCACCGTGGGCACACGACGACGAGAACCTCGCG 613  
 Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160  
 Db 614 CTGCATATCGTTGTTGGGACTCGCTGCCATGGATGTCTCTCAGTCCACCAACAGC 673  
 Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180  
 Db 674 ACATCTCTTGGCTACTCGGTGCTCGTGGGTACCGGCTTCTACGGGGACATCATCAAG 733  
 Qy 181 AspSerGluLysLysArgTyrPheLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200  
 Db 734 GACAGTGAGAAGAAACGGTGGTCTTGGCTTGGCAGATACGACTTTTCAGGTTTAAAGACC 793  
 Qy 201 PheLeuSerHisHisCysTyrGlyThrValSerPheLeuProAlaGlnHisThrVal 220  
 Db 794 TTCTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCTGCACACACACCGTG 853  
 Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240  
 Db 854 GGATCTCAAGGGATAGAGAGCCCTGCGGGCAGGATGCTTTGTTTGCAGCAAGCAAG 913  
 Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260  
 Db 914 CAGCAGCTGGAGGAGGAGCAGAGAAAGCACTGTATGGTTTGGAAAGCTGCGGAGCGTG 973  
 Qy 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280  
 Db 974 GAGGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGCTCTGT 1033  
 Qy 281 AlaCysArgArgSerProArgGlyLysSerProAlaAlaHisLeuGlyAspGlySerSer 300  
 Db 1034 GCTTGTGCGGGAGCCCGGAGGCTCTCCCGGCTGCCCATCTTGGGAGACGGGTCTTCT 1093  
 Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320  
 Db 1094 GACCTCATCTCATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAC 1153  
 Qy 321 ThrAsnGlnAspGln 326  
 Db 1154 ACCAACCCAGCAGGACCCAG 1171  
 RESULT 6  
 ABL40828  
 ID ABL40828 standard; cDNA; 4413 BP.  
 AC ABL40828;  
 XX  
 XX 03-JUL-2002 (first entry)  
 XX  
 XX Human sphingosine kinase-like protein encoding cDNA.  
 XX  
 XX Human sphingosine kinase-like protein; intracellular signalling; gene;  
 XX cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;  
 XX autoimmune disease; rheumatoid arthritis; Parkinson's disease; ss.  
 XX  
 XX Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 XX CDS 1..1689  
 XX FT /\*tag= a  
 XX FT /product= "sphingosine kinase-like protein"  
 XX FT /note= "see ABB07857"  
 XX FT 76..1689  
 XX FT /\*tag= b  
 XX FT /product= "sphingosine kinase-like protein"







XX 29-OCT-2002; 2002US-0422472P.  
XX (GETH ) GENENTECH INC.  
XX Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;  
XX Wood WI, Wu TD;  
XX WPI; 2004-376182/35.  
XX P-PSDB; ADP55248.  
XX  
XX New PRO polynucleotides and polypeptides, useful in diagnosing  
XX and treating an immune related disease, e.g. systemic lupus  
XX erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in  
XX stimulating an immune response.  
XX  
XX Claim 2; SEQ ID NO 1223; 3009bp; English.  
XX  
XX The present invention describes an isolated PRO nucleic acid (I). Also  
XX described: (1) a vector comprising (1); (2) a host cell comprising the  
XX vector of (1); (3) a process for producing a PRO polypeptides; (4) an  
XX isolated PRO polypeptide; (5) a chimeric molecule comprising the  
XX polypeptide of (4) fused to a heterologous amino acid sequence; (6) an  
XX antibody which specifically binds to a polypeptide of (4); (7) a  
XX composition of matter comprising a polypeptide of (4), an agonist or  
XX antagonist of the polypeptide or an antibody that binds to the  
XX polypeptide in combination with a carrier; (8) an article of manufacture  
XX comprising a container, a label on the container and a composition of  
XX matter of (7); (9) a method of treating an immune related disease in a  
XX mammal; (10) a method for determining the presence of a PRO polypeptide  
XX in a sample suspected of having the polypeptide; (11) a method of  
XX diagnosing an immune related disease or an inflammatory immune response  
XX in mammal; (12) a method of identifying a compound that inhibits or  
XX mimics the activity of or expression of a gene encoding a PRO polypeptide  
XX ; and (13) a method of stimulating the immune response in a mammal. The  
XX PRO sequences have anti-allergic, anti-nausea, anti-arthritis,  
XX anti-asthmatic, anti-diabetic, anti-inflammatory, anti-psoriatic,  
XX anti-rheumatic, antithyroid, CNS, dermatological, gastrointestinal,  
XX haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,  
XX nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and  
XX virucide activities, and can be used in gene therapy. The nucleic acid  
XX (I) and the encoded polypeptides, compositions, kits and methods are  
XX useful in diagnosing and treating an immune related disease and in  
XX stimulating an immune response. The present sequence represents a human  
XX PRO nucleotide sequence from the present invention.  
XX  
XX SQ Sequence 4445 BP; 1018 A; 1134 C; 1216 G; 1077 T; 0 U; 0 Other;  
XX  
XX Alignment Scores:  
XX Pred. No.: 4,97e-170 Length: 4445  
XX Score: 1640.50 Matches: 314  
XX Percent Similarity: 96.63% Conservative: 1  
XX Best Local Similarity: 96.32% Mismatches: 0  
XX Query Match: 95.54% Indels: 11  
XX DB: 13 Gaps: 1  
XX  
XX US-10-631-958-2 (1-326) x ADP55247 (1-4445)  
XX  
XX QY 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20  
XX DB 511 CCNAGCAATTACTGTTATATATCAACCCGTTTGAGAGAAAGCAAGGACGCGGATA 570  
XX  
XX QY 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40  
XX DB 571 TATGAAAGAAAGTGGCCACCACTGTTCACTTAGCCTCCTCATCCACCTGACATCATC--- 627  
XX  
XX QY 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60  
XX DB 628 -----GTTACTGAACATGCTTAATCAGGCCAAGGAG 657  
XX  
XX QY 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80  
XX DB 658 ACTCTGTATGAGATTACATAGACAAATACGACGCGCATCGTGTGTGCGGAGATGGT 717

QY 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100  
DB 718 ATGTTTACGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGAGCGCGGGTTCGAC 777  
QY 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120  
DB 778 CAGAACCAACCCCGGGCTGTGTCCTCCAGTAGCTCCGATTCGGATTTGGAAATCATNTCCCGCA 837  
QY 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140  
DB 838 GGGTCAACGGACTGCTGTGTTACTCCACCGTGGGACACGACGACGAGAACCTCGCGG 897  
QY 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160  
DB 898 CTGCATATCGTGTGTTGGGAGCTCGCTGGCCATGGATGTGCTCAGTCCACCAACAGC 957  
QY 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180  
DB 958 ACATCTCTCCCACTGCTATGAAGGACAGTGTCTCTCTCCCTGTCACACACACACG 1017  
QY 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200  
DB 1018 CACAGTGAGAGAAACGGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACC 1077  
QY 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220  
DB 1078 TTCTCTCTCCCACTGCTATGAAGGACAGTGTCTCTCTCTCCCTGTCACACACACACG 1137  
QY 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240  
DB 1138 GGATCTCCAAAGGATAGGAAGCCCTGCCGGGACAGATGCTTTGTTGAGGCAAGCAAG 1197  
QY 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260  
DB 1198 CAGCAGCTGGAGGAGGAGGACAGAGAACACTGTATGGTTTGGAAAGCTGCGGAGGACGTG 1257  
QY 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280  
DB 1258 GAGGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCCACCAACATGTCTCTGT 1317  
QY 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300  
DB 1318 GCTTGTGCGCGAGGCCCGGAGGCTCTCCCGGCTGCCACCTGGGGAGACGGGTCTTCT 1377  
QY 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320  
DB 1378 GACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAC 1437  
QY 321 ThrAsnGlnGlnAspGln 326  
DB 1438 ACCAACCAAGCAGGACCCAG 1455  
XX  
XX RESULT 9  
XX ABX70921  
XX ID ABX70921 standard; cDNA; 4432 BP.  
XX AC ABX70921;  
XX XX  
XX DT 05-MAR-2003 (first entry)  
XX XX  
XX DE Novel human cDNA sequence #146.  
XX XX  
XX KW Human; gene; ss; nervous system disorder; peripheral neuropathy;  
XX KW Huntington's disease; amyotrophic lateral sclerosis; haemophilia;  
XX KW neurodegenerative disease; Parkinson's disease; Alzheimer's disease;  
XX KW autoimmune disease; systemic lupus erythematosus; rheumatoid arthritis;  
XX KW insulin-dependent diabetes mellitus; anaemia; thrombocytopenia; wound;  
XX KW ulcer; burn; bone disorder; osteoporosis; osteoarthritis; stroke;  
XX KW fibrosis; reperfusion injury; infection; allergic rhinitis; asthma;  
XX KW coagulation disorder; cancer; tumour; inflammatory disease; septic shock;  
XX KW Crohn's disease; anaphylaxis; proliferation; chemotactic; chemokinetic;  
XX KW differentiation; stem cell growth factor; haematopoiesis; chemokinetic;

haemostatic; antiinflammatory; expressed sequence tag; EST.

KW OS Homo sapiens.  
 XX WO200281731-A2.  
 XX 17-OCT-2002.  
 XX 29-JAN-2002; 2002WO-US001222.  
 XX 30-JAN-2001; 2001US-00774528.  
 XX (HYSE-) HYSEQ INC.  
 XX (GOOD/) GOODRICH R W.  
 XX Tang TY, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
 XX Xue AJ, Yang Y, Wehrman T, Wang J, Wang D, Drmanac RT;  
 XX WPI; 2003-058563/05.  
 XX Novel polypeptide useful for treating neurodegenerative diseases, myeloid  
 XX or lymphoid cell disorders, bone disorders, mechanical and traumatic  
 XX disorders, coagulation disorders, and inflammatory diseases.  
 XX Claim 1; Page: 612pp; English.  
 XX This invention relates to the cDNA sequences encoding an isolated novel  
 XX human polypeptide. The protein encoded by the nucleic acid of the  
 XX invention is useful for treating central and peripheral nervous system  
 XX diseases (e.g. peripheral neuropathy, Huntington's disease, amyotrophic  
 XX lateral sclerosis); neurodegenerative diseases (e.g. Parkinson's disease,  
 XX Alzheimer's disease); autoimmune disease (e.g. systemic lupus  
 XX erythematosus, rheumatoid arthritis, insulin-dependent diabetes mellitus)  
 XX ; myeloid or lymphoid cell disorders (e.g. anaemia and thrombocytopaenia)  
 XX ; wounds, ulcers, burns; bone disorders (e.g. osteoporosis,  
 XX osteoarthritis); mechanical and traumatic disorders (e.g. stroke, head  
 XX trauma); lung or liver fibrosis; reperfusion injury in various tissues;  
 XX bacterial, viral or fungal infections; allergic conditions such as  
 XX allergic rhinitis, asthma; coagulation disorders (e.g. haemophilia);  
 XX cancer and tumours; and inflammatory diseases (e.g. septic shock, Crohn's  
 XX disease, anaphylaxis). The protein may be used to inhibit the growth,  
 XX infection or function of infectious agents such as bacteria, fungi,  
 XX viruses, or to effect bodily characteristics, biorhythms or circadian  
 XX cycles of rhythms. The protein may also have  
 XX proliferation/differentiation, stem cell growth factor, haematopoiesis  
 XX regulation, immune stimulating or suppressing, chemotactic/chemokinetic,  
 XX haemostatic and thrombolytic, receptor/ligand, and antiinflammatory  
 XX activities. The cDNA sequences of the invention are useful for expressing  
 XX recombinant protein for analysis. The present sequence represents a novel  
 XX human cDNA sequence of the invention, this sequence is an expressed  
 XX sequence tag (EST) and was identified using subtractive hybridisation  
 XX  
 SQ Sequence 4432 BP; 970 A; 1122 C; 1282 G; 1058 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2,946-169 Length: 4432  
 Score: 1633.50 MatChes: 313  
 Percent Similarity: 96.32% Conservative: 1  
 Best Local Similarity: 96.01% Mismatches: 1  
 Query Match: 95.14% Indels: 11  
 DB: 8 Gaps: 1  
 US-10-631-958-2 (1-326) x ABX70921 (1-4432)  
 QY 1 ProlysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20  
 DB 1591 CCAAAGCATTTTACTGGTATTTATCAACCCGTTTGGAGGAAAAGGCAAGGCAAGCGGATA 1650  
 QY 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40  
 DB 1651 TATGAAGAGAAAGTGGCCACTGTTTCACCTTAGCTTCATCACCACCTGACATCATC--- 1707  
 QY 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60

1708 -----GTTACTGAACATGCTAATCAGGCAAGAG 1737  
 QY 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80  
 DB 1738 ACTCTGATGAGATTAACTAGACAAATACAGCGCATCTCTGTGTGGCGGAGATGGT 1797  
 QY 81 MetPheSerGluValLeuHisGlyIleGlyArgThrGlnArgSerAlaGlyValAsp 100  
 DB 1798 ATGTTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGAGCGCGGGGTGCAC 1857  
 QY 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120  
 DB 1858 CAGAACCAACCCCGGGCTGCTGGTCCCCAGTAGCTCCGGATTGGAATCATTTCCGCA 1917  
 QY 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140  
 DB 1918 GGGTCAACGGACTGCTGTGTACTCCACCGTGGGACACGACGACGAGAAACCTCGCG 1977  
 QY 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160  
 DB 1978 CTGCATATCGTTGTTGGGACTCGCTGCCCATGGATGTCTCCTCAGTCCACCAACAGC 2037  
 QY 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleIys 180  
 DB 2038 ACATCTCTCGCTACTCTCGTGTCTGCTGGCTTACGGCTTCTACGGGGACATCATCAAG 2097  
 QY 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200  
 DB 2098 GACAGTGAGAAGAAACCGTGTGGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGACC 2157  
 QY 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220  
 DB 2158 TTCTCTCTCCACCACTGCTATGGAGGGACAGTGTCTTCTCTCCCTGCACAAACACG 2217  
 QY 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerIys 240  
 DB 2218 GGATCTCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTGCGAGCAAGCAAG 2277  
 QY 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260  
 DB 2278 CAGCAGCTGGAGGAGGAGCAGAGAAGAACACTGATGTGTTGGAACTGCGGAGACGTG 2337  
 QY 261 GluIleTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280  
 DB 2338 GAGCAGTGGCAAGTCGTCGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGCTCTGT 2397  
 QY 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300  
 DB 2398 GCTTGTGCGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCT 2457  
 QY 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320  
 DB 2458 GACCTCATCTCTATCCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAC 2517  
 QY 321 ThrAsnGlnGlnAspGln 326  
 DB 2518 ACCAACCCAGCAGGACCAG 2535  
 RESULT 10  
 ABA96945  
 ID ABA96945 standard; cDNA; 4463 BP.  
 XX  
 AC ABA96945;  
 XX  
 DT 20-MAY-2002 (first entry)  
 XX  
 DE Human ceramide kinase hCERK1-encoding cDNA.  
 XX  
 KW Human; ceramide kinase; hCERK1; drug screening; gene therapy;  
 KW neurological disease; inflammation; human immunodeficiency virus;  
 KW HIV infection; type 2 diabetes; obesity; sepsis; arteriosclerosis;  
 KW cancer; neuroprotective; antiinflammatory; anti-HIV; antidiabetic;

KW anorectic; antibacterial; antiseptic; antiarteriosclerotic; cytostatic;  
 XX gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 124..1737

XX /tag= a

XX /product= "Human ceramide kinase hCERK1"

XX WO200196575-A1.

XX 20-DEC-2001.

XX 11-JUN-2001; 2001WO-JP004889.

XX 14-JUN-2000; 2000JP-00178039.

XX (SANY ) SANKYO CO LTD.

XX Sugiura M, Kono K, Kohama T;

XX WPI; 2002-179513/23.

XX P-PSDB; AAM49115.

XX Human ceramide kinase gene and the enzyme encoded by it for screening  
 PT substances as drugs for neurological, inflammatory and other disorders.

XX Claim 5; Page 46-53; 61pp; Japanese.

CC This sequence represents cDNA encoding a human ceramide kinase designated  
 CC hCERK1. The invention relates to hCERK1, nucleic acids encoding it,  
 CC expression vectors and host cells containing hCERK1 nucleic acids, the  
 CC recombinant production of hCERK1 and antibodies specific for hCERK1. The  
 CC invention also encompasses methods of isolating hCERK1 from samples, the  
 CC use of hCERK1 in drug screening, and the use of hCERK1 nucleic acid  
 CC sequences in gene therapy. hCERK1 mediates the ATP-dependent 1-  
 CC phosphorylation of ceramides and can be used to screen for therapeutic  
 CC and preventive agents for a wide range of disorders. Such disorders  
 CC include neurological disease, inflammation, human immunodeficiency virus  
 CC (HIV) infection, type 2 diabetes, obesity, sepsis, arteriosclerosis and  
 CC cancer

SQ Sequence 4463 BP; 1026 A; 1135 C; 1217 G; 1084 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 3,83e-169 Length: 4463  
 Score: 1632.50 Matches: 312  
 Percent Similarity: 96.32% Conservative: 2  
 Best Local Similarity: 95.71% Mismatches: 1  
 Query Match: 95.08% Indels: 11  
 DB: 6 Gaps: 1

US-10-631-958-2 (1-326) x ABA96945 (1-4463)

Qy 1 ProLysHisLeuLeuValPheIleAsnProPheGlyLysGlyGlnGlyLysArgIle 20  
 Db 511 CCAAGCATTACTGTTATTTATCAACCGTTTGGAGAAAAGGACAAAGCGGATA 570  
 Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40  
 Db 571 TATGAAGAAGAAGTGCCACCACTGTTACCTTAGCCTCCATCACCACCTGACATCATC--- 627  
 Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60  
 Db 628 -----GTTACTGAACATGCTTAATCAGGCCAAGGAG 657  
 Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyLysGly 80  
 Db 658 ACTCTGTATGAGATTAAATACATACAAATACGACGCGCATCGTGTGTCGGCGGAGATGT 717  
 Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100

Db 718 ATGTTACGGAGGTCTCTCACGGTCTGATTGGGAGGACGACAGAGCGCGGGGTGCAC 777  
 Qy 101 GlnAenHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120  
 Db 778 CAGAACCAACCCCGGGCTGTGCTGGTCCCGAGTAGCCTCCGATTGGAATCATTCGCCGA 837  
 Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140  
 Db 838 GGTCCAACGGACTGCGTGTGTACTCCACCGTGGGCACACGACGACGAGAACCTCGCGG 897  
 Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160  
 Db 898 CTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCTAGTCCACCAACACAGC 957  
 Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180  
 Db 958 ACACCTCTTCGCTACTCCGCTGTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAG 1017  
 Qy 181 AspSerGluLysLysArgTyrPheLeuAlaArgTyrAspPheSerGlyLeuLysThr 200  
 Db 1018 GACAGTGAGAAAGAACCGTGGTCTTCCACGATACGACTTTTTCAGGTTTAAAGACC 1077  
 Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220  
 Db 1078 TTCTCTCTCCACCACTGCTATGAAGGGACAGTGTCTTCTCTCCCTGCAACACACACGGTG 1137  
 Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240  
 Db 1138 GGATCTCAAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTGGAGCAAGCAAG 1197  
 Qy 241 GlnGlnLeuGluGluGlnLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260  
 Db 1198 CAGCAGCTGGAGGAGGAGCAGAGAAGACACTGTATGTTTGGAACTGCGGAGGACGTG 1257  
 Qy 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280  
 Db 1258 GAGGAGTGGCAAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCAACAAACATGCTCTGT 1317  
 Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300  
 Db 1318 GCTTGTGCGGAGGCCGCCAGGGGCTCTCCCGGCTGCGCCTTGGGAGACGGGTCTTCT 1377  
 Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320  
 Db 1378 GACCTCATCTCATCCGGAATGCTCCAAGTTCAATTTCTGAGATTTCTCATCAGGCAC 1437  
 Qy 321 ThrAsnGlnAspGln 326  
 Db 1438 ACCAACCCAGCAGGAGCCAG 1455  
 RESULT 11  
 ID ADN62844  
 XX ADN62844 standard; DNA; 1740 BP.  
 AC ADN62844;  
 XX ADN62844;  
 DT 01-JUL-2004 (first entry)  
 XX Human NOV9a DNA.  
 DE ds; gene; human; NOVX; metabolic disorder; diabetes; obesity;  
 XX infectious disease; anorexia; cancer; cancer-associated cachexia;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; hematopoietic disorder; dyslipidaemia;  
 KW metabolic syndrome X; wasting disorder.  
 XX Homo sapiens.  
 OS US2004038223-A1.  
 PN 26-FEB-2004.  
 PD 01-OCT-2002; 2002US-00262511.  
 PF

```

XX 02-OCT-2001; 2001US-0326483P.
PR 05-OCT-2001; 2001US-0327435P.
PR 05-OCT-2001; 2001US-0327449P.
PR 09-OCT-2001; 2001US-0327917P.
PR 09-OCT-2001; 2001US-0328029P.
PR 09-OCT-2001; 2001US-0328044P.
PR 09-OCT-2001; 2001US-0328056P.
PR 12-OCT-2001; 2001US-0328849P.
PR 15-OCT-2001; 2001US-0329414P.
PR 17-OCT-2001; 2001US-0330142P.
PR 18-OCT-2001; 2001US-0330309P.
PR 22-OCT-2001; 2001US-0341058P.
PR 24-OCT-2001; 2001US-0339266P.
PR 24-OCT-2001; 2001US-0343629P.
PR 29-OCT-2001; 2001US-0349575P.
PR 01-NOV-2001; 2001US-0346357P.
PR 17-APR-2002; 2002US-0373260P.
PR 19-APR-2002; 2002US-0373815P.
PR 19-APR-2002; 2002US-0373817P.
PR 19-APR-2002; 2002US-0373826P.
PR 19-APR-2002; 2002US-0373884P.
PR 22-APR-2002; 2002US-0374977P.
PR 16-MAY-2002; 2002US-0381037P.
PR 16-MAY-2002; 2002US-0381038P.
PR 16-MAY-2002; 2002US-0381042P.
PR 17-MAY-2002; 2002US-0381042P.
PR 28-MAY-2002; 2002US-0383656P.
PR 29-MAY-2002; 2002US-0383831P.
PR 25-JUN-2002; 2002US-0391335P.
XX (SMIT/) SMITHSON G.
PA (MILL/) MILLET I.
PA (PEYM/) PEYMAN J A.
PA (KEKU/) KEKUDA R.
PA (JUJJ/) JU J.
PA (LILL/) LI L.
PA (GUOX/) GUO X.
PA (PATT/) PATTURAJAN M.
PA (SPYT/) SPYTEK K A.
PA (EDIN/) EDINGER S R.
PA (ELLE/) ELLERMAN K.
PA (MALY/) MALYANKAR U M.
PA (ORTT/) ORT T.
PA (GORM/) GORMAN L.
PA (ZERH/) ZERHUSEN B D.
PA (ANDE/) ANDERSON D W.
PA (ZHON/) ZHONG M.
PA (CATT/) CATTERTON E.
PA (JIWW/) JI W.
PA (MILL/) MILLER C E.
PA (RAST/) RASTELLI L.
PA (STON/) STONE D J.
PA (PENA/) PENA C E A.
PA (SHEN/) SHENOY S G.
PA (SHIM/) SHIMKETS R A.
PA (ROTH/) ROTHENBERG M E.
PA (LEAC/) LEACH M D.
PA (AGEE/) AGEE M L.
PA (BERG/) BERGHS C.
PA (DIPV/) DIPIPPO V A.
PA (EISE/) EISEN A.
PA (GANG/) GANGOLLI E A.
PA (RIEG/) RIEGER D K.
PA (SPAD/) SPADERNA S K.
XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;
PI Shmkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;
PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;
XX
DR WPI; 2004-213931/20.
DR P-PSDB; ADN62845.
XX Isolated NOVX polypeptides and nucleic acids, useful for preventing,
PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.
XX Claim 20; SEQ ID NO 39; 395pp; English.
XX The invention relates to isolated NOVX polypeptides and polynucleotides.
CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or
CC treat a medical condition in human related to the aberrant expression and
CC activity of NOVX polypeptides. For example, NOVX polypeptides and
CC polynucleotides may be used to treat disorders associated with decreased
CC expression or activity of NOVX by supplementing the patient our
CC production or to rectify mutations. Conversely, antisense NA molecules
CC may be administered to down regulate expression of NOVX polypeptides by
CC binding with the cells own genes and preventing their expression. NOVX
CC polynucleotides and complementary sequences may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar sequences in samples, and so which patients may be in need of
CC restorative therapy. NOVX polypeptides may also be used as antigens in
CC the production of antibodies and in assays to identify modulators
CC (agonists and antagonists) of the expression and activity of NOVX. The
CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be
CC used to modulate NOVX polynucleotide expression and activity of NOVX
CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as
CC diagnostic agents for detecting the presence of NOVX in samples. NOVX
CC polypeptides and polynucleotides may be used in this way to prevent,
CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious
CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative
CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,
CC haematopoietic disorders, and the various dyslipidaemias, metabolic
CC disturbances associated with obesity, the metabolic syndrome X and
CC wasting disorders associated with chronic diseases and various cancers.
CC They may also be used as antibacterial agents. The present sequence
CC represents DNA encoding a human NOVX protein.
XX
SQ Sequence 1740 BP; 388 A; 477 C; 530 G; 345 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.18e-165 Length: 1740
Score: 1595.50 Matches: 312
Percent Similarity: 96.02% Conservative: 2
Best Local Similarity: 95.41% Mismatches: 1
Query Match: 92.92% Indels: 13
DB: Gaps: 1

US-10-631-958-2 (1-326) x ADN62844 (1-1740)
Qy 1 ProlYshHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyeArgIle 20
Db 463 CCAAGCATTACTGGTATTATTCACCCGTTTGGAGGAAAAGGACAAGGCAAGCGGATA 522
Qy 21 TyrGluArgLysValAlaProLeuPheThrIleuAlaSerIleThrAspIleIleGly 40
Db 523 TATGAAAGAAAAGTGGCACCACCTGTTCCACTTACCTCCATCCACCTAGCATCATC--- 579
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 580 -----GTTACTGAACATGCTATATCAGCCCAAGAG 609
Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAsp-Gl 80
Db 610 ACTCTGTATGAGATTACATAGACAATACACGCGCAT-GTCTGTGTGCGCGGAGATCGG 668
Qy 80 yMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAs 100
Db 669 TATGTTTCAGCGAGTGTGTCACGCTCTGATTGGGAGGACGCGAGAGGCGCGCGGTCGA 728
Qy 100 pGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAl 120
Db 729 CCAGAACACCCCGGGCTGTGCTGGTCCCGAGTAGCCTCCGGATTGGATTCATCCCGC 788

```

QY 120 aGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAl 140  
Db 789 AGGGTCAACGGACTGGTGTGTACTCCACCGTGGGCACACGCGACGCAAAACCTCGGC 848  
QY 140 aLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSe 160  
Db 849 GCTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCTAGTCCACACACACG 908  
QY 160 rThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLy 180  
Db 909 CACACTCTCTCGTACTCTCGTGTCCCTGTGGCTACGGCTTCTACGGGGACATCAAC 968  
QY 180 sAspSerGluValValArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLyTh 200  
Db 969 GCACAGTGAAGAACAACGGTGGTGTCTCCAGATACGACTTTTCAGGTTTAAAGAC 1028  
QY 200 rPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVa 220  
Db 1029 CTTCTCTCCACCACTGCTATGAGGGACAGTGTCTCTCTCCCTGACACACACCGGT 1088  
QY 220 lGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLy 240  
Db 1089 GGGATCTCCAGGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTGCGGCAAGCAA 1148  
QY 240 sGlnGlnLeuGluGluGlnLysIleValLeuTyrGlyLeuGluAlaAlaGluaspVa 260  
Db 1149 GCAGCAGCTGGAGGAGGAGCAAGAAGACATGTATGGTTTGGAAAGCTGCGGAGCGT 1208  
QY 260 lGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCy 280  
Db 1209 GGAGGAGTGGCAAGTCTGTGTGGAGATTTCTGGCCATCAATGCCAACAACATGTCTCTG 1268  
QY 280 sAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSe 300  
Db 1269 TGCTTGTGCGCGAGGCCCGAGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTC 1328  
QY 300 rAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHi 320  
Db 1329 TGACCTCATCTCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 1388  
QY 320 sThrAsnGlnGlnAspGln 326  
Db 1389 CACCACCCAGCAGGACCAG 1407  
RESULT 12  
ID ADS10370  
XX ADS10370 standard; DNA; 4702 BP.  
XX AC ADS10370;  
XX DT 16-DEC-2004 (first entry)  
XX DE Human therapeutic DNA - SEQ ID 607.  
XX KW antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnery;  
KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;  
KW aplastic anaemia; cancer; wound healing; gene therapy; ds; gene.  
XX OS Homo sapiens.  
XX PN WO2004080148-A2.  
XX PD 23-SEP-2004.  
XX PF 30-SEP-2003; 2003WO-US030720.  
XX PR 02-OCT-2002; 2002US-0416186P.  
XX PA (NUVE-) NUVELO INC.  
XX PI Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y;  
PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;  
XX

DR WPI; 2004-668857/65.  
DR P-PSDB; ADS11054.  
XX New polynucleotide, useful in preparing a composition for diagnosing or  
PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,  
PT aplastic anemia or cancer for promoting wound healing.  
XX Claim 1; SEQ ID NO 607; 718pp; English.  
XX The invention relates to a novel isolated polynucleotide and the encoded  
CC polypeptide. The molecules of the invention demonstrate antiinflammatory,  
CC neuroprotective, antianaemic, cytostatic and vulnerary activities and may  
CC be useful in preparing a composition for diagnosing or treating  
CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell  
CC disorders, such as aplastic anaemia or cancer, as well as for promoting  
CC wound healing. The molecules may also be utilized during gene therapy  
CC procedures. The current sequence is that of a human therapeutic DNA of  
CC the invention. The current sequence is not shown explicitly within the  
CC specification but can be accessed from the WIPO web-site.  
XX  
SQ Sequence 4702 BP; 970 A; 1226 C; 1364 G; 1118 T; 0 U; 24 Other;  
Alignment Scores:  
Pred. No.: 4, 95e-145 Length: 4702  
Score: 1414.50 Matches: 276  
Percent Similarity: 80.58% Conservative: 2  
Best Local Similarity: 80.00% Mismatches: 0  
Query Match: 82.38% Indels: 67  
DB: 13 Gaps: 1  
US-10-631-958-2 (1-326) x ADS10370 (1-4702)  
QY 49 GluValIleThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAsp 68  
Db 1627 AAGCTGATTACTGAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAC 1686  
QY 69 LysTyrAspGlyIleValCysValGlyCysValGlyMetPheSerGluValLeuHisGly 88  
Db 1687 AAATACGACGCGACTGCTGTGTGCGCGAGATGGTATGTTCAGCGAGGTGTGACGGT 1746  
QY 89 LeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAenHisProArgAlaValLeu 108  
Db 1747 CTGATTGGGAGACGACGAGGAGCGCGGGTGCACCAAGAACCCCGGCTGTGCTG 1806  
QY 109 ValProSerSerLeuArgIleGlyIleIleProAla 120  
Db 1807 GTCCCCAGTAGCCTCCGGATTGGAATCATTCGCCGAGGCTTTGGCAAATGACCTGCGCAGG 1866  
QY 120 120  
Db 1867 GACGAGGTGTCTGTCTCTCTGCGCCCTGTGTCTGCGCCCGAGGGTGGCGCATGGTGCAC 1926  
QY 120 120  
Db 1927 ACTTTCACCTCGTTCACAGCTCTGGGATGTGAGCACCGCAGTCATCCCATTTTATGGAT 1986  
QY 121 121  
Db 1987 GAAGACAGGAGGACTGGGGACATGTGGCCCCCGGTGAGACGCTGGTGGCTTGGACGGG 2046  
QY 122 SerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeu 141  
Db 2047 TCAACGGACTCGTGTGTGTACTCCACCGTGGGCACCGACGACGAGAAACCTCGCGCTG 2106  
QY 142 HisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThr 161  
Db 2107 CATATCGTTGTGGGACTCGCTGGCCATGATGTCTCAGTCCACCAACAGCAC 2166  
QY 162 LeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAsp 181  
Db 2167 CTCCTTCGCTACTCGTGTCTCTGCTGGCTACGGCTTCTACGGGGACATCATCAAGGAC 2226  
QY 182 SerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPhe 201





QY 208 GluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLys 227  
 Db 1552 GAAGGACAGTGTCTCTCCCTGCACACACACCGTGGGATCTCCAGGAATAGGAAG 1611  
 QY 228 ProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLeuGluGln 247  
 Db 1612 CCTCGCGGAGGATGCTTTGTTGTCAGGCAAGCAGCAGCTGGAGGAGGAGCAG 1671  
 QY 248 LysLysAlaLeuTyrGlyLeuGluAlaGluAppValGluGluTrpGlnValValCys 267  
 Db 1672 AAGAAGACACTGTATGGTTTGAAGCTGCGGAGGACGTGGAGAGTGGCAAGTCTGT 1731  
 QY 268 GlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArg 287  
 Db 1732 GGAAGATTTCTGGCCATCAATGCCAACAATGTCGTGTGTGCGGAGCCCAAG 1791  
 QY 288 GlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIleLeuLys 307  
 Db 1792 GGCCTCTCCCGGCTGCCCACTTGGGAGAGCGGTCTTCTGACCTCATCTCATCCGAAA 1851  
 QY 308 CysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsnGlnGlnAspGln 326  
 Db 1852 TGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGCACACCAACAGCAGGACCAG 1908  
 RESULT 14  
 AAS77730  
 ID AAS77730 standard; cDNA; 2186 BP.  
 XX  
 AC AAS77730;  
 DT 13-FEB-2002 (first entry)  
 DE DNA encoding novel human diagnostic protein #13534.  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US008631.  
 XX  
 PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Drmanac RT, Liu C, Tang YT;  
 XX  
 DR WPI: 2001-639362/73.  
 DR P-PSDB; ABG13543.  
 XX

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

Claim 1; SEQ ID NO 13534; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 2186 BP; 455 A; 637 C; 569 G; 425 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.: 1.81e-135 Length: 2186  
 Score: 1323.50 Matches: 275  
 Percent Similarity: 58.26% Conservative: 0  
 Best Local Similarity: 58.26% Mismatches: 0  
 Query Match: 77.08% Indels: 197  
 DB: 5 Gaps: 3

US-10-631-958-2 (1-326) x AAS77730 (1-2186)

QY 52 ThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAsp 71  
 Db 3 ACTGAACATGCTAATCAGGCCAACGAGACTCTGTATGAGATTAAACATAGACAAATACGAC 62  
 QY 72 Gly----- 72  
 Db 63 GGATGCTACACTGCCATTCAAGGATGCCCTGTCTCGAAACCACAGCCCGCTGGGAAG 122  
 QY 72 ----- 72  
 Db 123 GCTGCCCGGTCAAGGTGTAGGCTACGGGGAAGGGCAAGAGCCTTCCCATGGAGCG 182  
 QY 72 ----- 72  
 Db 183 GGCCCAAGCAGATGCCTGTCACAGGGCCAGGTGTGCAGGTGCCCTCTCTGAAGCTGAA 242  
 QY 72 ----- 72  
 Db 243 GCCTCAGGCTGTGCGCTCAGAGGGCGGCACCTGCAGGAGCCCTGCAGCGCTGTGTCC 302  
 QY 73 -----IleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeu 89  
 Db 303 TCCTGCAGCATCGTCTGTGTCGGCGAGATGGTATGTTTCAGCGAGGTCTGCACGGTCTG 362  
 QY 90 IleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuVal 109  
 Db 363 ATGGGAGGACGCAGAGAGCGCGGGTCCACAGAACCCACCCCGGGCTGTGTGCTGTC 422  
 QY 110 ProSerSerLeuArgIleGlyIleIleProAla----- 120  
 Db 423 CCCAGTAGCCTCCGGATTGGAAATCATTTCCCGCAGGCTTTGCCAAATGACCTGGCAGGGGAC 482  
 QY 120 ----- 120  
 Db 483 GAGGTGTCTGTCTCTGCGCCCTGTGTCTGCGCCCGAGGGTGGCGCATGTGTGCACACT 542  
 QY 120 ----- 120  
 Db 543 TTCACCTCTCGTFCACAGCTCTGGGATGTGAGCACCGCAGTCATCCCATTTTATGATGAA 602  
 QY 121 -----GlySer 122  
 Db 603 GACAGGAGGACTGGGAGCATGTGGGCCCGGTGAGAACGCTGTGGTGTGGACGGGTCA 662  
 QY 123 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 142  
 Db 663 ACGGACTCGGTGTGTACTCCACCGTGGGCACACGACGACGAGAACCTTCGCGCTGCAT 722  
 QY 143 IleValVal----- 145



```
Db 723 ATCGTTGTTGCTGCTGCCCCGAGGCTCGGAACCAACCCGCGCTCCGCCATACTGCTGT 782
XX |||||
XX 145 ----- 145
XX
XX 783 GGCAGTGGGCGACGGGCGAGCTGTGCTGGGCTGCCAGTGTGTGAACGATGCTCTGTGAG 842
XX
XX 146 -----GlyAspSerLeuAlaMetAspValSer 154
XX |||||
Db 843 GCCTCGAGGCTTCAGTCCAGGATCCAGAGCCCCGGGAGCTCGCTGGCCATGGATGTGTCC 902
XX
Qy 155 SerValHisHisAenSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPhe 174
XX |||||
Db 903 TCAGTCCACCAACAACAGCACACTCTTCGCTACTCCGTGTCCCTGCTGGGCTACGGCTTC 962
XX
Qy 175 TyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaAaTqTyrAsp 194
XX |||||
Db 963 TACGGGGACATCATCAAGGACAGTGAGAAACCGGTGGTCTTGCCAGATACGAC 1022
XX
Qy 195 PheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeu 214
XX |||||
Db 1023 TTTTCAGGTTTAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTC 1082
XX
Qy 215 ProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPhe 234
XX |||||
Db 1083 CCTGCACAAACACCGGTGGGATCTCCAAGGATAGGAAGCCCTGCCGGCAGGATGCTTT 1142
XX
Qy 235 ValCysArgGlnSerLysGlnGlnLeuGluGlnLysLysAlaLeuTyrGlyLeu 254
XX |||||
Db 1143 GTTTGCAGGCAAGCAAGCACAGCTGGAGAGAGAGCAAGAAGCACTGTATGGTTG 1202
XX
Qy 255 GluAlaAlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsn 274
XX |||||
Db 1203 GAAGTCGGAGGAGCTGGAGAGTGGCAAGTCTGTGGGAAGTTCTGGCCATCAAT 1262
XX
Qy 275 AlaThrAenMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHis 294
XX |||||
Db 1263 GCCAAACATGTCTCTGTCTGTGCGCGAGCCCGAGGGCCCTCTCCCGGCTGCCAC 1322
XX
Qy 295 LeuGlyAspGlySerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeu 314
XX |||||
Db 1323 TTGGGAGACGGGTCTTCTGACCTCATCTATCCGGAATGCTCCAGGTTCAATTTTCTG 1382
XX
Qy 315 ArgPheLeuIleArgHisThrAsnGlnAspGln 326
XX |||||
Db 1383 AGATTCTCATCAGGCACACCAACAGCAGGACAG 1418
XX

RESULT 15
AAC76031
ID AAC76031 standard; cDNA; 753 BP.
XX
AC AAC76031;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF1586 polynucleotide sequence SEQ ID NO:3171.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antiporiatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.
XX
OS Homo sapiens.
```

```
XX
PN
XX WO200058473-A2.
XX
PD 05-OCT-2000.
XX
XX 31-MAR-2000; 2000WO-US008621.
XX
XX 31-MAR-1999; 99US-0127607P.
XX 02-APR-1999; 99US-0127636P.
XX 05-APR-1999; 99US-0127728P.
XX 30-MAR-2000; 2000US-00540763.
XX
XX (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
XX WPI; 2000-602362/57.
DR P-PSDB; AAB41822.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease.
XX
XX Claim 5; Page 2392; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnerary;
XX antiporiatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
XX anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
XX cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
XX dermatological; immunosuppressive; antirheumatic; antithyroid; antianemic. The
XX antiviral; antifungal; antirheumatic; antinflammatory; and antianemic. The
XX sequences can be used for determining the presence of or predisposition
XX to, or preventing or treating pathological conditions associated with an
XX ORFX-associated disorder. The nucleic acids can be used to express ORFX
XX proteins in gene therapy vectors. The proteins and nucleic acids may be
XX used to treat cancers, proliferative disorders, neurodegenerative
XX disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
XX diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
XX storage, systemic lupus erythematosus, severe combined immunodeficiency
XX (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
XX disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
XX cartilage damage, nocturnal haemoglobinuria, antinflammatory disease; to
XX enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX
XX SQ Sequence 753 BP; 165 A; 195 C; 207 G; 186 T; 0 U; 0 Other;
XX

Alignment Scores:
Pred. No.: 1.72e-106 Length: 753
Score: 1055.00 Matches: 199
Percent Similarity: 99.50% Conservative: 0
Best Local Similarity: 99.50% Mismatches: 1
Query Match: 61.44% Indels: 0
DB: 3 Gaps: 0

US-10-631-958-2 (1-326) x AAC76031 (1-753)
Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
|||
Db 133 GGGTCAACGGACTCGGTGTGTACTCCACCGTGGGACCGACGACGAGAAACCTCGGCG 192
|||
Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerValHisHisAsnSer 160
|||
Db 193 CTGCATATCGTGTGTGGGACTCGCTGGCCATGGATGTGTCTCAGTCCACCAACAGC 252
|||
Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
|||
Db 253 ACACCTCTCGTACTCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAG 312
|||
Qy 181 AspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
|||
Db 313 GACAGTGAGAAGAAACCGTGGTGGGCTTTGCCAGATACGACTTTTCAGGTTTAAAGACC 372
|||
```

```
Qy 201 PheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 373 TTCTCTCCACCCACTGCTATGAAGGGACAGTGTCTTCTCCCTGCACACACACGGTG 432
Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db 433 GGATCTCCAAAGGGATAGGAAGCCCTGCCGGGCAGGATGCTTTGTTTGACAGGCAAGCAAG 492
Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 493 CAGCAGCTGGAGGAGGAGCAGAGAAAGCACTGTATGGTTTGGAAAGCTCGGAGGATGTG 552
Qy 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 553 GAGGAGTGCAAGTCGTCTGTGGGAAGTTTCTGGCCCATCAATGCCACAAACATGTCCTGT 612
Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
Db 613 GCTTGTGCGCGGAGCCCCCAGGGGCTCTCCCCGGCTGCCCACTTGGGAGACGGGTCTTCT 672
Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
Db 673 GACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCTGGCAC 732
```

Search completed: September 6, 2005, 11:24:23  
Job time : 542.298 secs

**This Page Blank (uspto)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: September 5, 2005, 18:18:34 ; Search time 5901.4 Seconds  
(without alignments)  
10410.363 Million cell updates/sec

Title: US-10-631-958-9  
Perfect score: 1614  
Sequence: 1 atggggggcagggggcggc.....agccagactcacacagctga 1614

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

- 1: gb\_est1:\*
- 2: gb\_est2:\*
- 3: gb\_hic:\*
- 4: gb\_est3:\*
- 5: gb\_est4:\*
- 6: gb\_est5:\*
- 7: gb\_est6:\*
- 8: gb\_gss1:\*
- 9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1116.8	69.2	1824	3	AK042077	Mus muscu
2	1115.2	69.1	4248	3	AK052269	Mus muscu
3	861.8	53.4	1063	4	BM479389	AGENCOURT
4	769.4	47.7	1059	5	BQ054406	AGENCOURT
5	744.4	46.1	1078	5	BQ063738	AGENCOURT
6	739.6	45.8	797	7	CK000755	AGENCOURT
7	713	44.2	1047	5	BQ057191	AGENCOURT
8	704.4	43.6	713	7	CK000525	AGENCOURT
9	700	43.4	820	6	CD655311	AGENCOURT
10	684.4	42.4	732	7	CF135528	UI-HF-BNO
11	653	40.5	653	5	EX952302	DKF2p781L
12	649	40.2	661	7	CN296312	1700053321
13	579.4	35.9	581	5	BP224560	BP224560
14	579.2	35.9	584	5	BP310011	BP310011
15	570.4	35.3	584	5	BP309990	BP309990
16	563.8	34.9	573	7	CF138275	UI-HF-BNO
17	551.2	34.2	758	6	CB246749	UI-M-F10-
18	543.2	33.7	550	7	CF138634	UI-HF-BNO
19	535.8	33.2	541	7	CN296311	1700053321
20	504	31.2	582	5	BP274786	BP274786
21	491	30.4	812	7	CK603033	AGENCOURT
22	490.2	30.4	1017	5	BQ879739	AGENCOURT
23	477.2	29.6	488	5	BU430459	UI-HF-BNO
24	475	29.4	550	2	AW499858	UI-HF-BNO

25	464.4	28.8	474	2	AW503999	UI-HF-BNO
26	454	28.1	763	7	CK364327	AGENCOURT
27	442	27.4	449	2	AW503483	UI-HF-BNO
28	433.8	26.9	607	6	CB581157	AMGNNUC:N
29	429.2	26.6	653	6	BY754042	BY754042
30	428	26.5	688	7	CN296313	1700053326
31	427.2	26.5	608	6	CB580936	AMGNNUC:N
32	404.2	25.0	662	2	BB388667	BB388667
33	402.8	25.0	576	6	CB608292	AMGNNUC:N
34	394	24.4	656	6	BY739924	BY739924
35	382.2	23.6	540	6	CB612897	AMGNNUC:N
36	381.4	23.6	553	6	CA578972	K0731A08-
37	376.6	23.3	545	6	CB611947	AMGNNUC:N
38	367.4	22.8	376	5	BU430866	UI-HF-BNO
39	366.4	22.7	680	2	BB235873	BB235873
40	359.2	22.3	385	5	BU428707	UI-HF-BNO
41	358.8	22.2	779	5	BU265377	603374309
42	353.6	21.9	819	7	CN235868	CN235868
43	349	21.6	759	5	BU264162	603814452
44	348.8	21.6	359	5	BU428926	UI-HF-BNO
45	333.2	20.6	582	4	BM735016	MONOL_23_

ALIGNMENTS

RESULT 1	AK042077	AK042077	1824 bp	linear	HTC 03-APR-2004
LOCUS	Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched library, clone:A630056D11 product:DA59H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS) (FRAGMENT) homolog [Homo sapiens], full insert sequence.				
DEFINITION	Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched library, clone:A630056D11 product:DA59H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS) (FRAGMENT) homolog [Homo sapiens], full insert sequence.				
ACCESSION	AK042077				
VERSION	AK042077.1 GI:26334912				
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	1 Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	2 Carninci, P., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsu, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer				
AUTHORS	Genome Res. 10 (11), 1757-1771 (2000)				
TITLE	20530913				
JOURNAL	11076861				
MEDLINE	11076861				
PUBMED	11076861				
REFERENCE	4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)				
AUTHORS	FANTOM Consortium.				
TITLE	Functional annotation of a full-length mouse cDNA collection				
JOURNAL	Nature 409, 685-690 (2001)				
REFERENCE	5				

AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		
JOURNAL	Nature 420, 563-573 (2002)		
REFERENCE	6 (bases 1 to 1824)		
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imorani, K., Iehii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saichoh, R., Saichoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN) Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]		
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/WRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/ Location/Qualifiers		
FEATURES	<p>1..1824</p> <p>/organism="Mus musculus"</p> <p>/mol_type="mRNA"</p> <p>/strain="C57BL/6J"</p> <p>/db_xref="FANTOM.DB:A630056D11"</p> <p>/db_xref="taxon:10090"</p> <p>/clone="A630056D11"</p> <p>/tissue_type="thymus"</p> <p>/clone_libs="RIKEN full-length enriched mouse cDNA library"</p> <p>/dev_stages="3 days neonate"</p> <p>85..1680</p> <p>/note="unnamed protein product; DA59H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS) (FRAGMENT) homolog [Homo sapiens] (SPTR Q9UGES, evidence: FASTY, 86.5%ID, 100%length, match=945) putative"</p> <p>/codon_start=1</p> <p>/protein_id="BAC31157.1"</p> <p>/db_xref="GI:26334913"</p> <p>/translation="MGANGAAEPLHSLVWVKRRCAVLSLEPARALLRWRSPEPGPSA PGADARSLVIEIAVEEKDCEKHSRWHKMPFATVHRVKRVRHHRWKWAR TFWGADEQLCHLWQLRGLLESLSRPHLLVFINPGKGQGRKRYEKTVAFLFTL ASITTEIIITHANOKETLVEINTDVSVDGIVCVGDMGFSEVLHGVIGRTQOASGID PNHPRVAVPTSLRIGLIPAGSTDCVSVGTNDAAETLSALHIIIGSLAIDVSSVHY HNTLRYSLVSLIGFGFDGLIDSEKRWMLVRYDFSLGKTLFLSHQYVEGTLFLFA QHTVGSPPDNKPCRCAGCVCRSQKQLEEEKALYLENAEBEWEQVTCGKFLAIN ATNMSCAPRSPGGLSPFHLGDGSSDILIRKCSRNFRLRLIRHNOEQDFTFV EYVRVKFHFPTSKVEDNDNSKEQEKQKFGKIKDRPSCSASRSSWNGDGVMSHS PAIEVRVHQLVRLFPARGIEES"</p>		
CDS	<p>Query Match 69.2%; Score 1116.8; DB 3; Length 1824;</p> <p>Best Local Similarity 81.2%; Pred. No. 1.3e-279;</p> <p>Matches 1309; Conservative 0; Mismatches 302; Indels 1; Gaps 1;</p>		
ORIGIN	<p>1 ATGGGGCGCGAGCGGGCGGCGGAGCGCGTCAATCCGTGCTGGGTGAAGCAGCAGCGC 60</p> <p>85 ATGGGGCAATGGGGCGGCGGAGCGCGTCAATCCGTGCTGGGTGAAGCAGCAGCGC 144</p> <p>61 TGGCGCGTGAAGCTGGAGCGCGCGCGGCTCTGCTGGCTGGTGGCGGAGCGCGGGGCC 120</p> <p>145 TGTGCGCTCAGCTGGAGCGCGCGCGGCTCTGCTACGCTGGTGGCGGAGCGCGGAGCC 204</p> <p>121 GGAGCGCGCGCGCGCGCGCGGATGCTGCTGTGCTCTGTATCTGAGATCATCGCGTT 180</p> <p>205 GGGCCCTCCGCGCGCGGTGCGGATGCCCGCTCGGTACTGGTTCGGAGATCATCGCTTT 264</p> <p>181 GAGGAAACAGACGCTTACCGGGAACATCAAGGCGAGTGGAAAAATGGCAGAAAAAG 240</p> <p>265 GAGGAAACAGGACTCGGAGAAACACGCTCCAGTGGCGATGGCATAAATGAAAAAT 324</p> <p>241 CCTTACGCTTTACAGTTCACTGTGTAAAGAGACGAGCGGCAACCGCTGGAAGTGGCG 300</p> <p>325 CCGTTCCATTCAAGTCCACCGCTGTGAAGCGAGTACGACATCACCGCTGGAAGTGGCA 384</p> <p>301 CAGGTGACTTTCTGCTGCCAGAGGAGCGTGTCACTTGTGGCTGCAGACCGCTGGG 360</p> <p>385 CGGTGACCTTCTGGAGCGCGGAGCGAGCTGTGTCACTGTGGCTGCAGACCGCTCGT 444</p> <p>361 GAGATGCTGGAGAAAGCTGACGTCCAGACCAAGCATTTACTGGTATTTATCAACCCGTT 420</p> <p>445 GGGCTGCTGGAGAGCGCTGACTTCAAGACCGAGCATTTGCTGGTATTTATCAACCCTT 504</p> <p>421 GGAGGAAAGGACAAAGCAAGCGGATATATGAAAGAAAAGTGGCACCCTGTTACCTTA 480</p> <p>505 GGAGGAAAGGTGAGGCAAGCGCATCTATGAAAAAACAGTGGCGCTCTGTTTACCTTG 564</p> <p>481 GCCTCCATCACCCTGACATCATCGTTACTGACATCTAATCAGGCGCAAGGAGACTGTG 540</p> <p>565 GCTTCCATCATCTCGGAGATCATCATTTACAGAGATGCCAACCAAGCCAGGAGACTTTA 624</p> <p>541 TATGAGATTAACTAGACAAATACGCGCATCGTCTGTGTGGCGGAGATGTTATGTTTC 600</p> <p>625 TACGAGATCAACACAGACAGCTATGATGGCATCGTGTGCTAGTGGGAGCGGATGTTTC 684</p> <p>601 AGCGAGTGTGACAGCTGTGATTTGGAGAGACGAGAGAGCGCGCGGGTGCAGCAGAAC 660</p> <p>685 AGCGAGTGTGATCGGCTGATTTGGAGGAGCGAGCAGAGAGCGCTGTGATCGACCCCAAT 744</p> <p>661 CACCCCGGCGTGTGCTGCCAGTAGCTCCCGATTGGAATCAATCCCGCAGGCTCA 720</p> <p>745 CACCCCGAGCGGTGCTGGTGGCCAGTACCTTCAGGATCGGCATCATACCGCAGGGTCC 804</p> <p>721 ACGAGCTCGCTGTGTTACTCCACCGTGGGACACGAGCGACGAGAAAACCTCGGCGCTGCAT 780</p> <p>805 ACAGATTGTGTGTTACTCAACAGTGGGCACAAACGACGAGAGACATCGGCTTTGCAC 864</p> <p>781 ATCGTTCTGGGACTCGCTGGCGCATGGATGTGCTCAGTTCACCAACAACAGCAGCACTC 840</p> <p>865 ATCATTTGGGACTCACTGGCAATAGAGCTGCTCTGTGCACTACTACCAACACGCTG 924</p> <p>841 CTTGCTACTCCGTGCTCCCTGCTGGCTAGGCTTTCTACGGGAGACATCATCAAGCAGCT 900</p> <p>925 CTGCGGTACTCGGTTTCTCTGCTGGGTACGGTTTCTACGGGGACTTAATCAAGAGACAGT 984</p> <p>901 GAGAAGAAAACGTTGGTGGTCTTGGCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTC 960</p> <p>985 GAAAAGAAAACGTTGGATGGGCTCTGTCGGGTATGATTTCTCAGGTTTGAAGACCTTCTCTC 1044</p> <p>961 TCCACCACTGATGAAGGAGAGTGTCTTCTCTCCCTGCAACAACACAGCGTGGATCT 1020</p> <p>1045 TCTCATGACTATGAAGGAGACACTGTCTCTCTCCACGACACAGCAGCAGCGTGGATCT 1104</p> <p>1021 CCAGGGATAGGAGAGCTCCGGGAGGATGCTTTGTTTTCAGGCAAGCAAGCAGCAG 1080</p> <p>1105 CCACGGGACAATAAACCTTCGCGGCTGGGTGCTTCTGTGTGAGGAGCAAGCAACAG 1164</p>		

Qy	1081	CTGGAGGAGGACAGAGAAAGCACTGTATGTGTTTGAAGCTCGGAGGAGCATGGGAGGAG	1144
Db	1165	CTGGAAGAAGAAGAGAGAAAGCCCTGTATGGCTTGGAGAACGCCGAGGAAATGGAAGAG	1224
Qy	1141	TGGCAAGTCGTCTCTGGGAAGTTTCTGGCCATCAATGCCCAACAACATGTCTCTGTGCTTGT	1200
Db	1225	TGGCAAGTGACATGTGGGAAGTTCTTGCCCATCAATGCCCAACAACATGTCTGTGCTTGT	1284
Qy	1201	CGCCGAGCCCCCAGGGGCCCTCTCCCCGGCTGCCACATCTGGGAGACGGGTCTTCTGACCTC	1260
Db	1285	CCTCGGAGCCCTGGGGGCCCTGCCCATTTGCCATCTGGGAGATGGTCTTCTGACCTC	1344
Qy	1261	ATCCTCATCCGGAATGCTCAGGTTCAATTTCTTGAGATTTCTCATCAGGCAACCAAC	1320
Db	1345	ATCCTTATCCGGAAGTCTCAGGTTCAACTTCTTGAGATTTCTCATCAGGCAACCAAC	1404
Qy	1321	CAGCAGACCAAGTTTGACTTTCACATTTTGTGAGTTTATCGCTCAAGAAATTCAGTTT	1380
Db	1405	CAGGAGGACCAAGTTCGACTTCACTTTTGTAAGTTTATCGAGTCAAGAAATTCACCTC	1464
Qy	1381	ACGTTCGAAGCACATGGAGGATGAGGACAGCGGACCTCAAGGAGGGGGGGAAGAGCGCTTT	1440
Db	1465	ACGTTCGAAGCACGTTGGAAGACGAGGACAATGATCTCGAAGGACCAAGAGAGCAGAAGTTT	1524
Qy	1441	GGGCACATTTTGACGACGACACCCCTCTCTGTCTGTCACCGTCTCCAAACAGCTCCTGGAAC	1500
Db	1525	GGGAAGATCTGCAAGGACAGACCCCTCTTGCACTTGTCTCAGCTCCAGAAAGCTCTCGAAC	1584
Qy	1501	TGGCAGGGGAGGTCTTCGACAGCCCTGCCATCGAGGTCAGAGTCAGACTCCGACGTGGTT	1560
Db	1585	TGCGATGGCGAAGTCATGCACAGCCCGGCCCATTTGAGGTCAGGTCACCTGCCAGCTGGTG	1644
Qy	1561	CGACTCTTTTCGACGAGGAATTTGAAG-AGAAATCCGAAGCCAGACTTCACACAGC	1611
Db	1645	CGCCTCTTTCGTCGGGAATCGAGGAAGATCATAGCAGAGACCCCAAGC	1696
RESULT 2			
AK052269			
LOCUS			
DEFINITION			
	AK052269	4248 bp	linear
	Mus musculus	13 days embryo heart cDNA, RIKEN full-length enriched library, clone:1330016D08 product:DA59H8.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS) (FRAGMENT) homolog [Homo sapiens], full insert sequence.	HTC 03-APR-2004
ACCESSION	AK052269		
VERSION	AK052269.1	GI:26342491	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus		
ORGANISM	Mus musculus (house mouse)		
REFERENCE			
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		
REFERENCE	3		
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitesunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		

**TITLE**  
**JOURNAL**  
**MEDLINE**  
**PUBMED**  
**REFERENCE**  
**AUTHORS**

RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

**TITLE**  
**JOURNAL**  
**REFERENCE**  
**AUTHORS**

The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)

**TITLE**  
**JOURNAL**  
**REFERENCE**  
**AUTHORS**

The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 4248)

**TITLE**  
**JOURNAL**  
**REFERENCE**  
**AUTHORS**

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,  
Takeda, Y., Tanaka, T., Tonaru, A., Toya, T., Yasunishi, A.,  
Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,  
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,  
Fax:81-45-503-9216)

**TITLE**  
**JOURNAL**  
**COMMENT**

cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL:http://genome.gsc.riken.jp/  
URL:http://fantom.gsc.riken.jp/  
Location/Qualifiers  
1. .4248  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:D330016D08"  
/db\_xref="taxon:10090"  
/clone="D330016D08"  
/tissue\_type="heart"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="13 days embryo"  
260\_1855  
/note="unmaned protein product; DA59H18.2 (NOVEL PROTEIN  
SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED  
PROTEINS) (FRAGMENT) homolog [Homo sapiens] (SPTP1Q9UGES,  
evidence: FASTY, 86.5%ID, 100%length, match=945)  
putative"  
/codon\_start=1  
/protein\_id="BAC34908.1"  
/db\_xref="GI:26342492"  
/translation="MGANGAAEPLSHVLMYKRRRCRAVSLPARALLRWMRSPEPGPGA  
PGADARSLVLSIAVEEDCEKSHKRWKMPNFATVHRVKRVRHHRKWKARV  
TPWDAEQLCHLWLOTLRGLLESSTRPKHLVFINPFGSGGQGRVYKTVKAPLFTL  
ASTFWILLITEHANQAKETLYELNTSDYGVICVGGDGMPSEVLHGVIGRTQSGAGID  
PNHRPLVPSTIRIGLIPAGSDTDCVYSTVGTNDASTSAHLIIGDLSAIDVSSVH  
HNTLRVLSVLGIVGYDGLIKDSEKRWMLGVRYDFSGLTFLSHQYVYEGTSLSPFA  
QTVGSPRDNKPCACGVCROKQLEEBEKKALYGLNAEEMEEQVTCGKFLAIN  
ATNMSCACSPGGLSPFAHLDGGSDLLIRKCSRPNFLRLIRHTNNOBDOPTFTV

**FEATURES**  
**source**

**CDS**

EVYRVKXPHFTSKHVEYEDNDNSKEQKQKFKICKDRPSTCTCSASRSSWNCDEVMHS  
PAIEVRVHCOLVRLFARGIEES"

ORIGIN

Query Match 69.1%; Score 1115.2; DB 3; Length 4248;  
Best Local Similarity 81.1%; Pred. No. 4.2e-279;  
Matches 1308; Conservative 0; Mismatches 303; Indels 1; Gaps 1;

Qy	1	ATGGGGGCGAGCGGGGCGGAGCCGCTGCAATCCGTCGTGCTGGTGAAGCAGCAGCGC	60
Db			
Db	260	ATGGGGGCAATGGGGGCGGAGCGCGCTGCATCCGTCGTGGTGAACGGCGAGCG	319
Qy	61	TGCGCGGTGAGCTGGAGCCGCGGGGCTCTGCTGCGCTGGTGGCGAGCCCGGGGCC	120
Db			
Db	320	TGTGCGGTGAGCTGGAGCCGCGGGGCTCTGCTACGCTGGTGGCGAGCCCGAGGCC	379
Qy	121	GGAGCGGCGCCCGGGCGGAGTGCCTGCTCTGTGCTGTATCTGAGATCATCGCGGTT	180
Db			
Db	380	GGGCGCTCGGCGCGGGTGGCGATGCCCGCTCGGTACTGGTGTGGAGATCATCGCTGT	439
Qy	181	GAGGAAACAGAGCTTCACGGGAAACATCAAGGCAGTGGAAATGCGAGAAATGGAAG	240
Db			
Db	440	GAGGAAAGACGACTGCGAGAAACACGCCCTCCAGTGGCCGATGGCAATAATGAAAT	499
Qy	241	CTTTACGCTTTTACAGTTTCACTGTGTAAGAGAGCAGCAGCGCACCGCTGGAAGTGGCG	300
Db			
Db	500	CCGTTGCAATTCACAGTCCACCGTGTGAAGCAGTACGACATCACCGCTGGAAGTGGCA	559
Qy	301	CAGTGACTTCTGCTGTCAGAGAGAGAGCTGTGTCACTTGTGGTGGCAGACCCCTGGG	360
Db			
Db	560	CGGGTGACTTCTGGAGCGCGAGCAGCTGTGTCACTGTGGTGGCAGACCCCTCGT	619
Qy	361	GAGTGTGGAGAGCTGACGCTCCAGACCAAGCAATTTACTGCTATTTATCAACCCGTTT	420
Db			
Db	620	GGCTGCTGGAGAGCTGACTTCAAGACCGAAGCACTTGGCTGGTATTCATCAACCCCTTC	679
Qy	421	GGAGAAAGAGCAAGGCAAGCGGATATATGAAAGAAAGTGGCACCCTGTTTACCTTTA	480
Db			
Db	680	GGAGGAAGGTCAAGGCAAGCGCATCTATGAAAGAAACAGTGGCGCTCTGTTTACCTTG	739
Qy	481	GCCTCATCAGCATGCAATCATGTTACTGAAACATGCTTAATCAGGCGCAAGGAGACTCTG	540
Db			
Db	740	CTTCCATCACTACGGAGATCATATTACAGAGCATGCCAACCAAGCCAGGAGAGACTTTA	799
Qy	541	TATGAGTTAATAGACAAATAGACGGCATCTGCTGTGCGGAGAGTGGTATGTTTC	600
Db			
Db	800	TACGAGATCAACACAGACAGCTATGATGGCATCTGTGCTGAGTGGGAGCGGATGTTTC	859
Qy	601	AGCGAGTGTCTCACGCTCTGATTGGGAGGACGAGAGAGCGCGGGGTCCAGCAGAAC	660
Db			
Db	860	AGCGAGTGTCTGCATGGGTGATTGGGAGGACGAGCAGAGCGGTGTTATCGACCCCAAT	919
Qy	661	CACCCCGGGTGTGCTGGTCCAGTAGCTTCCGGAATGGAATCATTTCCGCGAGGTTCA	720
Db			
Db	920	CACCCCGAGCGTGTGCTGGTCCAGTAGTACCTCAGGATCGGCATCATACCCGCGAGGTC	979
Qy	721	ACGAGTCCGTGTGTTACTCCACCGTGGGACCAAGCAGCAGAAACCTCGGCGCTGCAT	780
Db			
Db	980	ACAGATTGTGTGTTACTCAACAGTGGGCAACACGACGAGAGACATCGGCTTTGGAC	1039
Qy	781	ATCGTTGTTGGGACTCGCTGCGCATGATGTGCTCCTCAGTCCACCAACACAGCAGCTC	840
Db			
Db	1040	ATCATATTGGGACTCACTGGCAATAGACGTGCTCTGTGCACTACCAATAACACGCTG	1099
Qy	841	CTTCGCTACTCGTGTCCCTGCTGGGCTACGGCTTCTACGGGAGCATCATCAAGACAGT	900
Db			
Db	1100	CTCGGCTACTCGGTTTCTCTGCTGGGCTACCGTTTCTACGGGAGCTTAATCAAGACAGT	1159
Qy	901	GAGAGAAACGCTGCTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTC	960
Db			
Db	1160	GAAAGAAACGCTGATGGGCTCTGCTCGGTATGATTTTCTCAGGTTTGAAGACCTTCTC	1219
Qy	961	TCCACCACTGCTATGAAGGAGCAGTGTCTTCTCTCTCTGCAACACACAGGTGGGATCT	1020

Db	1220	TCTCATCAGTACTATGAAGGACACATGCTCTCTCTCCAGCAGCAGCGTGGGATCT	1279
Qy	1021	CCAGGGATAGGAAGCCCTCCGCGAGGATGCTTTGTTGTCAGGCAAGCAAGCAGCAG	1080
Db			
Db	1280	CCAGGACAAATAAACCTCCGCGGCTGGGTGCTTCTGTGTCAGGAGAGCAAGCAACAG	1339
Qy	1081	CTGAGAGAGCAGAGAAAGACATGATGCTTTGGAAGCTGCGAGGACGTCGGAGGAG	1140
Db			
Db	1340	CTGGAAGAAGAGAAAGCCCTGTATGGCTGGAGAACGCCGAGGAATGGAAGAG	1399
Qy	1141	TGCAAGTCTGTGTGGAAAGTTTCGGCCCATCAATGCAACCAATGCTCTGTGCTGT	1200
Db			
Db	1400	TGGCAAGTGACATGTGGAAAGTTCTTGCCCATCAATGCCACCAACATGCTGTGCTGT	1459
Qy	1201	GGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTCTGACCTC	1260
Db			
Db	1460	CCTCGGAGCCCTGGGGGCTGTCCCCATTTGCCCATCTGGGAGATGGGTCTTCTGACCTC	1519
Qy	1261	ATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTTCTCATCAGGACACCAAC	1320
Db			
Db	1520	ATCTTATCCGGAAGTCTCCAGGTTCACTTCTGAGATTTCTCATCCGACACGAC	1579
Qy	1321	CAGCAGCACGAGTTGACTTCACTTTGTTGAAATTTATTCGCTCAAGAAATTCAGTTT	1380
Db			
Db	1580	CAGGAGCACGAGTTGACTTCACTTTGTTGAAATTTATCGAGTCAAGAAATTCACCTTC	1639
Qy	1381	ACCTCGAAGCACATGGAGGATGAGGACAGCGACTCAAGGAGGGGGAGAGCGCTTT	1440
Db			
Db	1640	ACGTCGAAGCAGCTGGAAATACGAGGACAAATGACTCGAAGGAAACAAGAGAGCAGAGTTT	1699
Qy	1441	GGGCACATTTGCAGCAGCACCCCTCTCTGCTGTCACCCCTCTCCACAGCTCTCTGGAAC	1500
Db			
Db	1700	GGGAGATCTGCAAGGACAGACCTCTTGCACTTGCTCAGCTCCAGAGCTCTCTGGAAC	1759
Qy	1501	TGCGACGGGAGGTCTTGCAAGCCCTGCAAGGTCATCAGGTCAGAGTCTCAGCTCGGTT	1560
Db			
Db	1760	TGCGATCGCAAGTCTATGCAAGCCGCCATTTGAGTCAAGGTCCCACTGCGAGCTGGTG	1819
Qy	1561	CGACTCTTCCAGCAGGAATTTGAG-AGAAATCCGAGCCAGACTCACACAGC	1611
Db			
Db	1820	CGCTCTTTTCTCGGGAAATCGAGGAAGAGTATAGCAAGAACCCCAAGC	1871

RESULT 3

BM479389  
LOCUS  
DEFINITION AGENCOURT\_6418742 NIH\_MGC\_67 Homo sapiens cdna clone IMAGE:5502281  
5', mRNA sequence.

ACCESSION  
BM479389

KEYWORDS  
EST.

SOURCE  
Homo sapiens (human)

ORGANISM

Homo sapiens  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
1 (bases 1 to 1063)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM12140 row: c column: 18

High quality sequence stop: 665.

Location/Qualifiers

1. .1063

FEATURES

source



```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5502281"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC 67"
/notes="Organ: eye; Vector: pCMV-SORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Library constructed by Life Technologies."
```

ORIGIN

Query Match	53.4%;	Score	861.8;	DB	4;	Length	1063;
Best Local Similarity	94.7%;	Pred. No.	3.5e-213;				
Matches	924;	Conservative	0;	Mismatches	42;	Indels	10;
						Gaps	3;
QY	463	GCACCACTGTTTCCACTTACCTTACCTCCATCACCACTGACATCATCGTTTACTGAAACATGCTTAAT	522				
Db	1	GCACCACTGTTTCCACTTACCTTACCTCCATCACCACTGACATCATCGTTTACTGAAACATGCTTAAT	60				
QY	523	CAGGCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGTC	582				
Db	61	CAGGCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGTC	120				
QY	583	GGCGAGATGTTTTCAGCGAGTGTCTGACGGTCTGATTGGGAGACGACGAGGAGC	642				
Db	121	GGCGAGATGTTTTCAGCGAGTGTCTGACGGTCTGATTGGGAGACGACGAGGAGC	180				
QY	643	GGCGGGTGCACGACCAACCCCGGCTGTGCTGCCAGTAGCTCCGGATTGGA	702				
Db	181	GGCGGGTGCACGACCAACCCCGGCTGTGCTGCCAGTAGCTCCGGATTGGA	240				
QY	703	ATCATTCGCGAGGTCACGAGCTGCGTGTGTTACTCCACGTTGGGACACGACGACGA	762				
Db	241	ATCATTCGCGAGGTCACGAGCTGCGTGTGTTACTCCACGTTGGGACACGACGACGA	300				
QY	763	GAACTCTGGCGCTGCATATCGTTTGGGGACTCGCTGGCCATGGATGTGTCCTCAGTC	822				
Db	301	GAACTCTGGCGCTGCATATCGTTTGGGGACTCGCTGGCCATGGATGTGTCCTCAGTC	360				
QY	823	CACCACACAGCACACTCTTCGCTACTCCGTGTCCTGCTGGGCTACGGCTTCTACGGG	882				
Db	361	CACCACACAGCACACTCTTCGCTACTCCGTGTCCTGCTGGGCTACGGCTTCTACGGG	420				
QY	883	GACATCATCAAGACAGTGAGAAGAAACGGTGGTGGTCTTTCAGATACGACTTTTCA	942				
Db	421	GACATCATCAAGACAGTGAGAAGAAACGGTGGTGGTCTTTCAGATACGACTTTTCA	480				
QY	943	GGTTTAAAGACCTTCTCTCCACCACTGTATAGAGGACAGTGTCTTCTCCTCCTGCA	1002				
Db	481	GGTTTAAAGACCTTCTCTCCACCACTGTATAGAGGACAGTGTCTTCTCCTCCTGCA	540				
QY	1003	CAACACAGGTGGGATCTCCAGGNTAGGAAGCCCTCCCGGACAGGATGCTTGTGTC	1062				
Db	541	CAACACAGGTGGGATCTCCAGGNTAGGAAGCCCTCCCGGACAGGATGCTTGTGTC	600				
QY	1063	AGGCAAGCAAGCAGCAGCTGGAGGAGACAGAGAAAGCACTGTATGGTTTGAAGCT	1122				
Db	601	AGGCAAGCAAGCAGCAGCTGGAGGAGACAGAGAAAGCACTGTATGGTTTGAAGCT	660				
QY	1123	GGGAGGACGTGGAGAGTGGCAAGTGTCTGTGGGAAATTTCTGGCCATCAATGCCACA	1182				
Db	661	GGGAGGACGTGGAGAGTGGCAAGTGTCTGTGGGAAATTTCTGGCCATCAATGCCACA	720				
QY	1183	ACATGTCCTGTGTGTGCGCGAGCCCCCAGGGGCTCTCCCGGCTGCCACTTGGGA	1242				
Db	721	ACATGTCCTGTGTGTGCGCGAGCCCCCAGGGGCTCTCCCGGCTGCCACTTGGGA	780				
QY	1243	GACGGGTCTTTCAGCTCATCTCATCCGGAATGCTCCAGGTTCAAATTTTCTGAGATT	1302				
Db	781	GACGGGTCTTTCAGCTCATCTCATCCGGAATGCTCCAGGTTCAAATTTTCTGAGATT	840				

QY	1303	CTCATCGGCACACCAACCAAGCAGCAGCAGTTTGACTTCAC--TTTTGTTGAAGTTTATC	1360				
Db	841	CTCATCGGCACACCAACCAAGCAGCAGCAGTTTGACTTCAC--TTTTGTTGAAGTTTATC	900				
QY	1361	CGCTCAAG--AAATTCCAGTTTACGTGGAACACATCGGAGGATG-----AGGACAGCGA	1412				
Db	901	CCCGTCAGGAAATCCCAAGTTTACTCTCAAAACCCCTTGGGAGGATGAAGCAGCAGCGA	960				
QY	1413	CCTCAAGGAGGGGGG 1428					
Db	961	CCTCAAGGAGGGGGG 976					
RESULT 4							
BQ054406							
LOCUS							
DEFINITION	BQ054406	1059 bp	mRNA	linear	EST	29-MAR-2002	
	AGENCOURT 6771237 NIH_MGC_99	Homo sapiens	cDNA clone	IMAGE:5803668			
	5', mRNA sequence.						
ACCESSION	BQ054406						
VERSION	BQ054406.1	GI:19813746					
KEYWORDS	EST.						
SOURCE	Homo sapiens (human)						
ORGANISM	Homo sapiens						
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
REFERENCE	1 (bases 1 to 1059)						
AUTHORS	NIH-MGC http://mgi.nci.nih.gov/.						
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)						
JOURNAL	Unpublished (1999)						
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Lou Staudt cDNA Library Preparation: Rubin Laboratory DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LUCM2040 row: a column: 13 High quality sequence stop: 601.						
FEATURES							
	Location/Qualifiers						
	1..1059						
	/organism="Homo sapiens"						
	/mol_type="mRNA"						
	/db_xref="taxon:9606"						
	/clone="IMAGE:5803668"						
	/tissue_type="lymphoma, cell line"						
	/lab_host="DH10B (phage-resistant)"						
	/clone_lib="NIH_MGC_99"						
	/note="Organ: lymph; Vector: pORB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."						
ORIGIN							
Query Match	47.7%;	Score	769.4;	DB	5;	Length	1059;
Best Local Similarity	99.1%;	Pred. No.	4.2e-189;				
Matches	784;	Conservative	0;	Mismatches	6;	Indels	1;
						Gaps	1;
QY	395	ATTTACTGTTATTTATCAACCCGTTTGGAGAAAGGACGAGCGGATATATGAAA	454				
Db	13	ATTTACTGTTATTTATCAACCCGTTTGGAGAAAGGACGAGCGGATATATGAAA	72				
QY	455	GAAGAAGTGGCAACCACTGTTTCACTTAGCCTCCATCCACCTGACATCATCGTTACTGAAC	514				
Db	73	GAAGAAGTGGCAACCACTGTTTCACTTAGCCTCCATCCACCTGACATCATCGTTACTGAAC	132				
QY	515	ATGTAATCAGGCCAAGGAGACTCTGTATGAGATTACATAGACAAATACGACGCGCATCG	574				

Db 133 ATGCTAATCAGCCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGACGCATCG 192  
Qy TCTGTGTCGGCGGAGATGTTATGTTTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGC 634  
Db TCTGTGTCGGCGGAGATGTTATGTTTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGC 252  
Qy AGAGGAGCGCGGGGTTCGACAGAAACCAACCCCGGGGTGTCTGTGTCGCCAGTAGCCTCC 694  
Db AGAGGAGCGCGGGGTTCGACAGAAACCAACCCCGGGGTGTCTGTGTCGCCAGTAGCCTCC 312  
Qy GGATTGGAATCATTTCCCGCAGGTTCAACGGAATCGGTGTGTGTTTATCTCAACCGTGGGACCA 754  
Db GGATTGGAATCATTTCCCGCAGGTTCAACGGAATCGGTGTGTGTTTATCTCAACCGTGGGACCA 372  
Qy GCGAGCGAGAACTCCGCGCTGCTATCGTTGTTGGGACTCGCTGGCCATGATGTGT 814  
Db GCGAGCGAGAACTCCGCGCTGCTATCGTTGTTGGGACTCGCTGGCCATGATGTGT 432  
Qy CCTCAGTCCACCAACAGCACACTCCCTTCGCTACTCCGTGTCCTGCTGGGCTACGGCT 874  
Db CCTCAGTCCACCAACAGCACACTCCCTTCGCTACTCCGTGTCCTGCTGGGCTACGGCT 492  
Qy TCTACGGGACATCATCAAGGACAGTGAAGAAACCGTGTGTTGGGTCTTCCAGAGATCG 934  
Db TCTACGGGACATCATCAAGGACAGTGAAGAAACCGTGTGTTGGGTCTTCCAGAGATCG 552  
Qy ACTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCC 994  
Db ACTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCC 612  
Qy TCCCTGCAACAACACCGTGGGATCTCCAGGGATAGGAAGCCCTGCGGCGAGGATGCT 1054  
Db TCCCTGCAACAACACCGTGGGATCTCCAGGGATAGGAAGCCCTGCGGCGAGGATGCT 672  
Qy TTGTTTGAGCAAGCAAGCAGCAGCTGGAAGGAGGAGCAGAAAGCACTGTATGGGT 732  
Db TTGTTTGAGCAAGCAAGCAGCAGCTGGAAGGAGGAGCAGAAAGCACTGTATGGGT 732  
Qy TGGAGCTGCGGAGGACGTGGAGGAGTGGCAAGTCTGTGGG-AGATTTCGGCCATC 1173  
Db TGGAGCTGCGGAGGACGTGGAGGAGTGGCAAGTCTGTGGG-AGATTTCGGCCATC 792  
Qy 1174 AATGCCACAA 1184  
Db 793 AATGCCACAA 803

RESULT 5  
BQ063738  
LOCUS AGENCOURT\_6873251 NIH\_MGC\_99 Homo sapiens cDNA clone IMAGE:5925382  
DEFINITION 5', mRNA sequence.  
ACCESSION BQ063738  
VERSION BQ063738.1 GI:19891754  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1078)  
NIH-MGC <http://mgi.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Lou Staudt  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LCM2096 row: p column: 23  
High quality sequence stop: 640.  
Location/Qualifiers  
1. 1078  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5925382"  
/tissue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_99"  
/note="Organ: lymph; Vector: pOTB7; Site: 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Size-selected >500bp for average insert size  
1.8kb. Library constructed by Ling Hong in the laboratory  
of Gerald M. Rubin (University of California, Berkeley)  
using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
II RT (Life Technologies). Note: this is a NIH\_MGC  
Library."

ORIGIN

Query Match 46.1%; Score 744.4; DB 5; Length 1078;  
Best Local Similarity 97.5%; Pred. No. 1.4e-182;  
Matches 778; Conservative 0; Mismatches 16; Indels 4; Gaps 2;

Qy 268 AAGAGACACGACGCGCACCGCTGGAAGTGGCGCAGGTGACTTTCTGTGTCTCCAGAGGAG 327  
Db 1 AAGAGACACGACGCGCACCGCTGGAAGTGGCGCAGGTGACTTTCTGTGTCTCCAGAGGAG 60  
Qy 328 CAGCTGTGTCACTTGTGGCTGCAGACCCCTCGGAGATGCTGGAGAAAGTGAAGTCAAG 387  
Db 61 CAGCTGTGTCACTTGTGGCTGCAGACCCCTCGGAGATGCTGGAGAAAGTGAAGTCAAG 120  
Qy 388 CCAAGCATTTTACTGGTATTATCAACCGCTTTGGAGAAAGGACAAAGCAAGCGGATA 447  
Db 121 CCAAGCATTTTACTGGTATTATCAACCGCTTTGGAGAAAGGACAAAGCAAGCGGATA 180  
Qy 448 TATGAAAGAAAGTGGCACCACTGTTTCACTTAGCTCCATCACCACCTGACATCATCGTT 507  
Db 181 TATGAAAGAAAGTGGCACCACTGTTTCACTTAGCTCCATCACCACCTGACATCATCGTT 240  
Qy 508 ACTGAAATCATCTAATCAGGCGCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGAC 567  
Db 241 ACTGAAATCATCTAATCAGGCGCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGAC 300  
Qy 568 GGCATCGTGTGTGGCGGAGATGTTATGTTACGAGGCTGCTGACGGTCTGATTGGG 627  
Db 301 GGCATCGTGTGTGGCGGAGATGTTATGTTACGAGGCTGCTGACGGTCTGATTGGG 360  
Qy 628 AGGACGACAGGAGCGCGGGGTGCACAGAAACCAACCCCGGGGTGTGCTGGTCCCCAGT 687  
Db 361 AGGACGACAGGAGCGCGGGGTGCACAGAAACCAACCCCGGGGTGTGCTGGTCCCCAGT 420  
Qy 688 AGCCTCCGGATTGGAATCATTTCCCGCAGGGTCAACGGACTGCGTGTGTTACTCCACCGTG 747  
Db 421 AGCCTCCGGATTGGAATCATTTCCCGCAGGGTCAACGGACTGCGTGTGTTACTCCACCGTG 480  
Qy 748 GGCACGACGACGCGAGAAACCTCGGCGCTGCATATCGTTTGGGGACTCGCTGGCCATG 807  
Db 481 GGCACGACGACGCGAGAAACCTCGGCGCTGCATATCGTTTGGGGACTCGCTGGCCATG 540  
Qy 808 GATGTGCTCTCAGTCCACCAACAGCACACTCTCTCGTACTCCGTGTCTCTGCTGGGC 867  
Db 541 GATGTGCTCTCAGTCCACCAACAGCACACTCTCTCGTACTCCGTGTCTCTGCTGGGC 600  
Qy 868 TACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAAACCGGTGGTGGGTCTTGCC 927  
Db 601 TACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAAACCGGTGGTGGGTCTTGCC 660  
Qy 928 AGATACGACTTTTCAAGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTG 987  
Db 661 AGATACGACTTTTCAAGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTG 720

QY 988 TCCTTCTCTCTGCAACACACAGG--TGGATCTCCAAGGA--TAGGAGCCCTGCGG 1043  
|||||  
Db 721 TCCTTCTCTCTGCAACACACACAGGTTGGGATCTCCAGGGAATAAGGAAGCCCTGCCC 780  
|||||  
QY 1044 GGCAGGATGCTTTGTTTG 1061  
|||  
Db 781 GGGCCAGGGATGCTTTTG 798  
|||

RESULT 6  
CK000755  
LOCUS CK000755 797 bp mRNA linear EST 26-NOV-2003  
DEFINITION AGENCOURT\_16363467 NIH\_MGC\_221 Homo sapiens cDNA clone  
IMAGE:30707875 5', mRNA sequence.

ACCESSION CK000755  
VERSION  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 797)

TITLE NIH-MGC <http://mgs.nci.nih.gov/>.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

CONTACT: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)

Tissue Procurement: James Martin, University of Iowa

cDNA Library Preparation: M. Bento Soares, University of Iowa

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: NDAM1073 row: c column: 20

High quality sequence stop: 656.

Location/Qualifiers

1..797

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:30707875"

/lab\_host="DH10B Tona"

/clone\_lib="NIH MGC 221"

/note="Organ: mixed; Vector: pYX-Asc; Site1: EcoRI;

Site 2: NotI; Library is oligo-dT primed and directionally

cloned. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated with

EcoR I adaptor, digested with Not I and then cloned

directionally into pYX-Aac vector. Average insert size

4-5kb. Adaptors 5' (AATTCGGACAGG)3' and 5'd

(CCTGTCGG)3'. 3' Linker sequence - GCGGCGCTGAGGCC T18.

Sequencing primers 3' end: T3 promoter primer 5'd

(ATATACCTCACTAAAGGGA)3'. 5' End: T7 promoter primer 5'd

(TAATACGACTCACTATAGG)3'. Library was constructed in the

laboratory of M. Bento Soares. Note: this is a NIH\_MGC

Library"

## ORIGIN

Query Match 45.8%; Score 739.6; DB 7; Length 797;  
Best Local Similarity 97.3%; Pred. No. 2.3e-181;  
Matches 751; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 458 AAGTGGCACCCTGTTACCTTAGCCTCCATCACCACCTGACATCATCTTACTGAACATG 517

Db 13 ACGAGGCACCCTGTTACCTTAGCCTCCATCACCACCTGACATCATCTTACTGAACATG 72

QY 518 CTAATCAGGCCCAAGGAGACTCTGTATGAGATTACATAGACAAATACAGCGCATCGTCT 577  
|||||  
Db 73 CTAATCAGGCCCAAGGAGACTCTGTATGAGATTACATAGACAAATACAGCGCATCGTCT 132  
|||||  
QY 578 GTGTCCGCGGAGATGTTATGTTTCAAGCAGAGGTGCTGCACGGTCTGATTGGGAGGACGAGA 637  
|||||  
Db 133 GTGTCCGCGGAGATGTTATGTTTCAAGCAGAGGTGCTGCACGGTCTGATTGGGAGGACGAGA 192  
|||||  
QY 638 GGAGCGCCGGGGTGCACAGAAACACCCCGGGGTGTGTGGTCCCAAGTAGCTTCGGGA 697  
|||||  
Db 193 GGAGCGCCGGGGTGCACAGAAACACCCCGGGGTGTGTGGTCCCAAGTAGCTTCGGGA 252  
|||||  
QY 698 TTGGAATCATTTCCCGCAGGGTCAACGGACTCGGTGTTGGGACTCGCTGGCCATGATGTCTCT 757  
|||||  
Db 253 TTGGAATCATTTCCCGCAGGGTCAACGGACTCGGTGTTGGGACTCGCTGGCCATGATGTCTCT 312  
|||||  
QY 758 ACGCAGAAACCTCGCGCTGCATATCGTTGTTGGGACTCGCTGGCCATGATGTCTCT 817  
|||||  
Db 313 ACGCAGAAACCTCGCGCTGCATATCGTTGTTGGGACTCGCTGGCCATGATGTCTCT 372  
|||||  
QY 818 CAGTCCACCAACACAGCACACTCTTCGCTACTCCGTGTCTCTGTGGGCTACGGCTTCT 877  
|||||  
Db 373 CAGTCCACCAACACAGCACACTCTTCGCTACTCCGTGTCTCTGTGGGCTACGGCTTCT 432  
|||||  
QY 878 ACGGGGACATCATCAAGGACAGTGAAGAACCGTGGTGGGCTTTGCCAGATACGACT 937  
|||||  
Db 433 ACGGGGACATCATCAAGGACAGTGAAGAACCGTGGTGGGCTTTGCCAGATACGACT 492  
|||||  
QY 938 TTTCCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCC 997  
|||||  
Db 493 TTTCCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCC 552  
|||||  
QY 998 CTGCACACACACACGGTGGGATCTCCAAAGGATAGAAAGCCCTGCCGGCAGATGCTTTG 1057  
|||||  
Db 553 CTGCACACACACACGGTGGGATCTCCAAAGGATAGAAAGCCCTGCCGGCAGATGCTTTG 612  
|||||  
QY 1058 TTTCCAGGCAAGCAAGCAGCAGTGGAGGAGGAGCAGAAAGACATGTATGGTTTGG 1117  
|||||  
Db 613 TTTCCAGGCAAGCAAGCAGCAGTGGAGGAGGAGCAGAAAGACATGTATGGTTTGG 672  
|||||  
QY 1118 AAGCTGCGGAGGACGTGGAGGAGTGGCAAGTCTGTGGGAAGTTTCTGCCCATCAATG 1177  
|||||  
Db 673 AAGCTGCGGAGGACGTGGAGGAGTGGCAAGTCTGTGGGAAGTTTCTGCCCATCAAT 732  
|||||  
QY 1178 CCACAAACATGTCTGTCTTGTCTCCCGGAGCCCGAGGGGCTCTCCCGGCG 1229  
|||||  
Db 733 GCACAAACATGTCTGTCTTGTCTCCCGGAGCCCGAGGGGCTCTCCCGGCTGC 784  
|||||

## RESULT 7

BQ057191

LOCUS

DEFINITION

5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1047)

TITLE NIH-MGC <http://mgs.nci.nih.gov/>.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

CONTACT: Robert Strausberg, Ph.D.

Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)

Tissue Procurement: Lou Staud

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

```

http://image.llnl.gov
Plate: LUCM2062 row: 1 column: 15
High quality sequence stop: 535.
Location/Qualifiers
1. .1047
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/Clones="IMAGE:5812382"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGACGACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
```

ORIGIN	Identity:
Query Match Best Local Similarity Matches 825; Conservative 8;	44.2%; Score 713; DB 5; Length 1047; 94.7%; Pred. No. 2.1e-174; 0; Mismatches 35; Indels 11; Gaps 8;
Qy	395 ATTTACTGGTATTTTATCAACCCGTTTGGAGGAAAGGACAAAGCGGATATATGAAA 454
Db	11 ATTTACTGGTATTTTATCAACCCGTTTGGAGGAAAGGACAAAGCGGATATATGAAA 70
Qy	455 GAAAAGTGGCCACCACTGTTACCTTAGCTCCATCAACAATGACATCATCGTTACTGAAC 514
Db	71 GAAAAGTGGCCACCACTGTTACCTTAGCTCCATCAACAATGACATCATCGTTACTGAAC 130
Qy	515 ATGCTAATCAGGCCAACAGGAGACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCG 574
Db	131 ATGCTAATCAGGCCAACAGGAGACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCG 190
Qy	575 TCTGTGTCCGCGGAGATGGTATGTTACGCGAGGTGCTGCACGGTCTGATTGGGAGGACGC 634
Db	191 TCTGTGTCCGCGGAGATGGTATGTTACGCGAGGTGCTGCACGGTCTGATTGGGAGGACGC 250
Qy	635 AGAGGAGCGCGGGGTGACACAGAAACACCCCGGGCTGTCTGGTCCCGAGTACGCTCC 694
Db	251 AGAGGAGCGCGGGGTGACACAGAAACACCCCGGGCTGTCTGGTCCCGAGTACGCTCC 310
Qy	695 GGATTGGAAATCATTTCCGCGCAGGGTCAAGGACTCGGTGTGTTACTCCACCGTGGGCACCA 754
Db	311 GGAATGGAAATCATTTCCGCGCAGGGTCAAGGACTCGGTGTGTTACTCCACCGTGGGCACCA 370
Qy	755 GCGACGCGAGAAACCTCCGGCGTGCATATCGTTGTTGGGGACTCCGCTGGCCATGGATGTGT 814
Db	371 GCGACGCGAGAAACCTCCGGCGTGCATATCGTTGTTGGGGACTCCGCTGGCCATGGATGTGT 430
Qy	815 CCTCAGTCCACCAACAACAGCACATCCTTTCGCTACTCCGTGTCCTGCTGGGCTACGGCT 874
Db	431 CCTCAGTCCACCAACAACAGCACATCCTTTCGCTACTCCGTGTCCTGCTGGGCTACGGCT 490
Qy	875 TCTACGGGGACATCATCAAGGACAGTGAAGAAACGGTGTGGGTCTTGGCCAGATACG 934
Db	491 TCTACGGGGACATCATCAAGGACAGTGAAGAAACGGTGTGGGTCTTGGCCAGATACG 550
Qy	935 ACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAA - GGGACAGTGTCTTTC 993
Db	551 ACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAAGGGGACAGTGTCTTTC 610
Qy	994 CTCCTCTGCACAAACACCGGTGGGATCT - CCAAGGGATAGGAAGCCCTG - CCGGGCAGGAT 1051
Db	611 CTCCTCTGCACAAACACCGGTGGGATCTCCCAAGGGATAGGAACCTCTCCCGGGCACGAT 670
Qy	1052 GCTTTTG - TTTTGACGGCAAGCAACGACAGCTGGAGGAGGACGAGAAAGC -- ACTGT 1108

D	b	671	GCTTTGTTTGGCGGAAACGAAGCAGCACACTGGAGGAGAGGCCAAGAAGAACCACTGTG	730
Q	y	1109	ATGCTTTGGAGCTGCGGAGGACGTCGGAGAGTGCGCAAGTCGTCT- GTGGGAACTTTCTG	1167
D	b	731	ATGCTTTGGAACTCTGCGGAAAGACGCGGACGAGTGGCAAATCTCTCTGGGGGGAGCTTTCTG	790
Q	y	1168	GCCATCAATGCCACAACATGTCTCT-GTGCCTGTGCGCGGAGCCCCAGGGGCTCT--C	1233
D	b	791	GCCATCATGCCCCAAACCTGTCTGTGTGTGTGTGCGCGGAACCCCAGGGGCTTCTCC	850
Q	y	1224	CCGCGCTGCCCACTCTGGGAGACGGCTTCTTCT	1254
D	b	851	CCGCGCTGCCCAATTGGGGAAGGGTCTCT	881
RESULT 8				
CK000525				
LOCUS				
DEFINITION AGENCOURT 16369000 NIH MGC 221 Homo sapiens CDNA clone				
EST IMAGE:30708597 5', mRNA sequence.				
ACCESSION CK000525				
VERSION CK000525.1 GI:38526559				
KEYWORDS EST.				
SOURCE Homo sapiens (human)				
ORGANISM Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
1 (bases 1 to 713)				
NIH-MGC http://mgc.nci.nih.gov/.				
National Institutes of Health, Mammalian Gene Collection (MGC)				
Unpublished (1999)				
Contact: Daniela S. Gerhard, Ph.D.				
Office of Cancer Genomics				
National Cancer Institute / NIH				
Bldg. 31 Rm10A07 Bethesda, MD 20892				
Email: cga@nci.nih.gov				
Tissue Procurement: James Martin, University of Iowa				
CDNA Library Preparation: M. Bento Soares, University of Iowa				
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)				
Clone distribution: Agencourt Bioscience Corporation				
Clone distribution: MGC clone distribution information can be				
found through the I.M.A.G.E. Consortium/LLNL at:				
http://image.llnl.gov				
Plate: NDAMI075 row: a column: 22				
High quality sequence stop: 689.				
Location/Qualifiers				
1. .713				
/organism="Homo sapiens"				
/mol_type="mRNA"				
/db_xref="taxon:9606"				
/clones="IMAGE:30708597"				
/lab_host="DHIOB Tona"				
/clone_lib="NIH_MGC_221"				
/note="Organ: mixed; Vector: pYX-Asc; Site:1: EcoRI;				
Site 2: NotI; Library is oligo-dT primed and directionally				
cloned Denatured RNA was size fractionated on a 1% agarose				
gel. First strand cDNA synthesis was primed with oligo-dT				
primer containing a Not I site. Double strand cDNA was				
size selected according to mRNA size fraction, ligated with				
EcoR I adaptor, digested with Not I and then cloned				
directionally into pYX-Asc vector. Average insert size				
(4-Skb. Adaptors 5' (AATTCGACAGG)3' and 5'd				
CTCTGCGCG)3'. 3' linker sequence - GCGGCGCTCAGAGCC T18.				
Sequencing primers 3'end: T3 promoter primer 5'd				
(ATAACCCCTCAATAGGA)3'. 5' End: T7 promoter primer 5'd				
(TTATACGACTCATAGG)3'. Library was constructed in the				
laboratory of M. Bento Soares. Note: this is a NIH MGC				
Library"				
FEATURES				
source				

ORIGIN	Library	Query Match	43.6%	Score 704.4	DB 7	Length 713
		Best Local Similarity	99.9%	Pred. No. 3.4e-172		
		Matches 705	Conservative	0	Mismatches 1	Indels 0
						Gaps 0

QY 828 CAACAGCACACTCTTCGCTACTCCGTGTCCTGCTGGCTACGGCTTCTACGGGGACAT 887  
Db 8 CAACAGCACACTCTTCGCTACTCCGTGTCCTGCTGGCTACGGCTTCTACGGGGACAT 67  
QY 888 CATCAAGGACAGTGAAGAAACGGTGGTGGCTTCTGCGACATACGACTTTTCAGGTTT 947  
Db 68 CATCAAGGACAGTGAAGAAACGGTGGTGGCTTCTGCGACATACGACTTTTCAGGTTT 127  
QY 948 AAGACCTTCTCTCCACCACTGTATGAAGGGACAGTGTCTTCTCTCCCTGCGCAACA 1007  
Db 128 AAGACCTTCTCTCCACCACTGTATGAAGGGACAGTGTCTTCTCTCCCTGCGCAACA 187  
QY 1008 CACGCTGGGATCTCCAGGGATAGAGCCCTGCGGGCAGAGTCTTGTGTTGAGGCA 1067  
Db 188 CACGCTGGGATCTCCAGGGATAGAGCCCTGCGGGCAGAGTCTTGTGTTGAGGCA 247  
QY 1068 AAGCAAGCAGCAGCTGGAGGAGCAGAGAAAGCACTGTATGTTTGGAGCTGCGGA 1127  
Db 248 AAGCAAGCAGCAGCTGGAGGAGCAGAGAAAGCACTGTATGTTTGGAGCTGCGGA 307  
QY 1128 GGACCTGGAGGAGTGGCAAGTCTGTGGGAAAGTCTTGGCCATCAATGCCACAAACAT 1187  
Db 308 GGACCTGGAGGAGTGGCAAGTCTGTGGGAAAGTCTTGGCCATCAATGCCACAAACAT 367  
QY 1188 GTCTGTGCTGTGCTGCGGAGCCCGAGGGCTCTCCCGGCTGCCACTTGGGAGCGG 1247  
Db 368 GTCTGTGCTGTGCTGCGGAGCCCGAGGGCTCTCCCGGCTGCCACTTGGGAGCGG 427  
QY 1248 GTCTTCTGACCTCATCTCCGGAATGCTCCAGGTTCAATTTCTGAGATTTCTCAT 1307  
Db 428 GTCTTCTGACCTCATCTCCGGAATGCTCCAGGTTCAATTTCTGAGATTTCTCAT 487  
QY 1308 CAGGCACACCAACAGCAGGACCTTGTGACTTCACTTTTGTGAAGTTTATCGGTCAA 1367  
Db 488 CAGGCACACCAACAGCAGGACCTTGTGACTTCACTTTTGTGAAGTTTATCGGTCAA 547  
QY 1368 GAAATTCAGTTTACGTGAAGCACATGAGAGATGAGACACGACCTCAAGGAGGGGG 1427  
Db 548 GAAATTCAGTTTACGTGAAGCACATGAGAGATGAGACACGACCTCAAGGAGGGGG 607  
QY 1428 GAAGAAGCGCTTGGGACATTTGAGCAGCAGCCCTCTGCTGCTGCTGCTGCTGCTCAA 1487  
Db 608 GAAGAAGCGCTTGGGACATTTGAGCAGCAGCCCTCTGCTGCTGCTGCTGCTGCTCAA 667  
QY 1488 CAGCTCTCGAAGTGGAGGGAGGTCTGACAGCCCTGCGATC 1533  
Db 668 CAGCTCTCGAAGTGGAGGGAGGTCTGACAGCTGCGATC 713

RESULT 9  
CD655311  
LOCUS  
DEFINITION AGENCOURT\_14552675 NIA Human H1 Embryonic Stem Cell cDNA Library  
(Long) Homo sapiens cDNA clone IMAGE:30426593 5', mRNA sequence.  
ACCESSION CD655311  
VERSION CD655311.1 GI:31895467  
KEYWORDS EST.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 820)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Irene Ginis and Mahendra Rao, NIA  
cDNA Library Preparation: Yulan Piao and Minoru Ko

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC c lone distribution information  
can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: NDAMS12 row: k column: 18  
High quality sequence stop: 673.  
Location/Qualifiers  
1. 820  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30426593"  
/tissue\_type="Embryonic Stem cells"  
/cell\_line="WA01"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"  
/note="Vector: pCMV-Sport6; Site\_1: NotI; Site\_2: SalI;  
This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01/H1 cell line. Undifferentiated human ES cell line WA01/H1 was obtained from WiCell Research Institute, Inc., Madison, WI, cultured according to their instructions, on MEF feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UTP, TERT, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1, TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days after plating), the ES cells from 4 X 6cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with Trizol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID:11544199]) Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTCTAGATCGGAGCGCCGCTTTTCTTTT-3'] from 3.4g of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lona-linker LL-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by Centricon-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S for 25 cycles. The products were purified by phenol/chloroform extraction and Centricon-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The average insert size is about 3.6kb."

ORIGIN

Query Match 43.4%; Score 700; DB 6; Length 820;  
Best Local Similarity 97.1%; Pred. No. 4.9e-171;  
Matches 712; Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
QY 738 CTCACCGTGGGACACGACGACGAGAAACCTCGGCGCTGCATATCGTTGTGGGACTC 797  
Db 11 CTCACCGTGGGACACGACGACGAGAAACCTCGGCGCTGCATATCGTTGTGGGACTC 70  
QY 798 GCTGGCCATGATGTGCTTACGGGGACATCATCAAGGACAGTGCAGAGAAACCGTGGTT 857  
Db 71 GCTGGCCATGATGTGCTTACGGGGACATCATCAAGGACAGTGCAGAGAAACCGTGGTT 130  
QY 858 CCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAGGACAGTGCAGAGAAACCGTGGTT 917  
Db 131 CCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAGGACAGTGCAGAGAAACCGTGGTT 190  
QY 918 GGGTCTTGGCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGA 977  
Db 191 GGGTCTTGGCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGA 250  
QY 978 AGGAGACAGTCTCTCTCTCTGACACACACCGTGGGATCTCCAGGGATAGGAAGCC 1037

Db 251 AGGACAGTGTCTTCTCCTCGACACACACCGTGGGATCTCAAAGGATAGGAAGCC 310  
Qy 1038 CTGCGGCGCAGGATGCTTTGTTTGAGCAAGCAAGCAGCAGCTGGAGGAGGAGCAGAA 1097  
Db 311 CTGCGGCGCAGGATGCTTTGTTTGAGCAAGCAAGCAGCAGCTGGAGGAGGAGCAGAA 370  
Qy 1098 GAAAGCACTGATGTTTGGAGCTGCGGAGCAGTGGAGGAGTGGCAAGTCGTCTGTGG 1157  
Db 371 GAAAGCACTGATGTTTGGAGCTGCGGAGCAGTGGAGGAGTGGCAAGTCGTCTGTGG 430  
Qy 1158 GAAATTTCTGCGCCATCAATGCAAAACATGTCCTGTGCTTGTGCGCGAGCCCGAGGG 1217  
Db 431 GAAATTTCTGCGCCATCAATGCAAAACATGTCCTGTGCTTGTGCGCGAGCCCGAGGG 490  
Qy 1218 CCTCTCCCGGCTGCGCCATGCGGAGCGGTCTTCTGACCTCATCTCATCCGGAATG 1277  
Db 491 CCTCTCCCGGCTGCGCCATGCGGAGCGGTCTTCTGACCTCATCTCATCCGGAATG 550  
Qy 1278 CTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAACCAGCAGGACCAAGTTGA 1337  
Db 551 CTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAACCAGCAGGACCAAGTTGA 610  
Qy 1338 CTTACATTTTGTGAAGTTATCGGTCAAGAAATTCAGTTTACGTGCAAGCAGCATGGA 1397  
Db 611 CTTACATTTTGTGAAGTTATCGGTCAAGAAATTCAGTTTACGTGCAAGCAGCATGGA 670  
Qy 1398 GGATGAGCAGCAGCAGCTCAAGGAGGGGGGAAGCGCTTGGGCAATTTTCAGCAG 1457  
Db 671 GGATGAGCAGCAGCAGCTCAAGGAGGGGGGAAGCGCTTGGGCAATTTTCAGCAGCAG 1457  
Qy 1458 CCACCCCTCTGC 1470  
Db 731 CCTCTGCTGCTGC 743

RESULT 10  
CF135528  
LOCUS  
DEFINITION  
IMAGE:3090211 5', mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: Louis Staudt  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
http://genome.uiowa.edu/distribution/humanfl.html  
Seg primer: pyx-5.  
Location/Qualifiers  
1. 732  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"

/clones="IMAGE:3090211"  
/tissue\_type="lymph"  
/cell\_type="germinal center B cells"  
/cell\_line="MGC85"  
/lab\_host="DH10B (LTI)"  
/clone\_lib="NIH\_MGC\_50"  
/notes="Vector: p773-Pac; Site 1: NotI; Site 2: Eco RI;  
Constructed from size fractionated cytoplasmic mRNA  
(3.5-4.4kb). Directionally cloned. Cells provided by  
Louis M. Staudt, Ph.D. Library preparation by Maria de  
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

ORIGIN

Query Match 42.4%; Score 684.4; DB 7; Length 732;  
Best Local Similarity 99.7%; Pred. No. 5.5e-167;  
Matches 696; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 918 GGGTCTTGGCAGATACGACTTTTCAAGCTTTAAAGACCTTCTCTCCACCACTGCTATGA 977  
Db 8 GGGTCTTGGCAGATACGACTTTTCAAGCTTTAAAGACCTTCTCTCCACCACTGCTATGA 67  
Qy 978 AGGACAGTGTCTTCTCCTCGACACACCGTGGGATCTCCAGGGATAGGAAGCC 1037  
Db 68 AGGACAGTGTCTTCTCCTCGACACACCGTGGGATCTCCAGGGATAGGAAGCC 127  
Qy 1038 CTGCGGCGCAGGATGCTTTGTTTGAGCAAGCAAGCAGCAGCTGGAGGAGGAGCAGAA 1097  
Db 128 CTGCGGCGCAGGATGCTTTGTTTGAGCAAGCAAGCAGCAGCTGGAGGAGGAGCAGAA 187  
Qy 1098 GAAAGCACTGATGTTTGGAGCTGCGGAGGAGCTGGAGGAGTGGCAAGTCGTCTGTGG 1157  
Db 188 GAAAGCACTGATGTTTGGAGCTGCGGAGGAGCTGGAGGAGTGGCAAGTCGTCTGTGG 247  
Qy 1158 GAAATTTCTGCGCCATCAATGCCAACAATGTCCTGTGCTTGTGCGCGAGCCCGAGGG 1217  
Db 248 GAAATTTCTGCGCCATCAATGCCAACAATGTCCTGTGCTTGTGCGCGAGCCCGAGGG 307  
Qy 1218 CCTCTCCCGGCTGCGCCATGCGGAGCGGTCTTCTGACCTCATCTCATCCGGAATG 1277  
Db 308 CCTCTCCCGGCTGCGCCATGCGGAGCGGTCTTCTGACCTCATCTCATCCGGAATG 367  
Qy 1278 CTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAACCAGCAGGACCAAGTTGA 1337  
Db 368 CTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAACCAGCAGGACCAAGTTGA 427  
Qy 1338 CTTACATTTTGTGAAGTTATCGGTCAAGAAATTCAGTTTACGTGCAAGCAGCATGGA 1397  
Db 428 CTTACATTTTGTGAAGTTATCGGTCAAGAAATTCAGTTTACGTGCAAGCAGCATGGA 487  
Qy 1398 GGATGAGCAGCAGCAGCTCAA- GGAGGGGGGGAAGAGCGCTTTGGGCACATTTTCAGCA 1456  
Db 488 GGATGAGCAGCAGCAGCTCAAAGGGAGGGGGGAAGAGCGCTTTGGGCACATTTTCAGCA 547  
Qy 1457 GCACACCTCTCTGCTGTGTCACCGTCTTCCAAACAGCTCTTGGAACTGGAGCGGAGGTCC 1516  
Db 548 GCCACCCCTCTCTGCTGTGTCACCGTCTTCCAAACAGCTCTTGGAACTGGAGCGGAGGTCC 607  
Qy 1517 TGACACCCCTGCGCATCGAGTTCAGCTCCACAGCTGGTTTCGACTCTTTTCACAGAG 1576  
Db 608 TGACACCCCTGCGCATCGAGTTCAGCTCCACAGCTGGTTTCGACTCTTTTCACAGAG 667  
Qy 1577 GAATTGAAGAGATCCCAAGCCAGACTCACACAGCTGA 1614  
Db 668 GGATTGAAGAGATCCCAAGCCAGACTCACACAGCTGA 705

RESULT 11  
BX952302  
LOCUS  
DEFINITION  
ACCSSION  
VERSION  
BX952302.1 GI:43428907  
BX952302 653 bp mRNA linear EST 01-MAR-2004  
DKFZp781l1183 r1 781 (synonym: hlcc4) Homo sapiens cDNA clone  
DKFZp781l1183 5', mRNA sequence.  
BX952302  
BX952302.1 GI:43428907



Db	481	GAATTTCCAGTTTACGTCGAAGCACATGGAGGATGAGGACAGCGACCTCAAGGAGGGGGG	54
Qy	1428	GAAGAGCGCTTTGGGCACATTTTGACAGAGCACCCTTCCTGCTGCTGACCGTCTCCAA	148
Db	541	GAAGAGCGCTTTGGGCACATTTTGACAGAGCACCCTTCCTGCTGCTGACCGTCTCCAA	600
Qy	1488	CAGCTCTCGAACTCGACGGGAGGTCTCTGACAGCCCTGCAATCGAGGTCA	1540
Db	601	CAGCTCTCGAACTCGACGGGAGGTCTCTGACAGCCCTGCAATCGAGGTCA	653
RESULT 12			
CN296312			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			

```
Db 241 GCAAGCAGCAGCTGGAGGAGGACAGAGAAGCACTGTATGTTTGGAAAGCTGCGGAGG 300
Qy 1130 ACETGGAGGAGTGGCAAGTCTGTGTGGAAAGTTTCTGGCCATCAATGCCCAACAACATGT 1189
Db 301 ACGTGGAGGAGTGGCAAGTCTGTGTGGAAAGTTTCTGGCCATCAATGCCCAACAACATGT 360
Qy 1190 CCTGTGCTTGTGCGCGGAGCCCGAGGGCCCTCTCCCGGCTGCCACTTGGGAGACGGGT 1249
Db 361 CCGTGTGCTTGTGCGCGGAGCCCGAGGGCCCTCTCCCGGCTGCCACTTGGGAGACGGGT 420
Qy 1250 CTTCTGACCTCATCTCATCCGGAATCTCCAGGTCAATTTCTGAGATTTCTCATCA 1309
Db 421 CTTCTGACCTCATCTCATCCGGAATCTCCAGGTCAATTTCTGAGATTTCTCATCA 480
Qy 1310 GGCACACCAACAGCAGGACCACTTGCATTTTGTGTAAGTTTATCGCTCAAGA 1369
Db 481 GGCACACCAACAGCAGGACCACTTGCATTTTGTGTAAGTTTATCGCTCAAGA 540
Qy 1370 AATTCCAGTTTACGTCGAAGCACATGGAGGATGAGGACGAGCCTCAAGGA-GGGGGGG 1428
Db 541 AATTCCAGTTTACGTCGAAGCACATGGAGGATGAGGACGAGCCTCAAGGAGGGGGGG 600
Qy 1429 AAGAAGCGCTTTGGGACATTTGAGCAGCAGCCCTCTCTGCTGTCGACCGTCTCCAAC 1488
Db 601 AAGAAGCGCTTTGGGACATTTGAGCAGCAGCCCTCTCTGCTGTCGACCGTCTCCAAC 660
Qy 1489 A 1489
Db 661 A 661

RESULT 13
BP224560 581 bp mRNA linear EST 15-SEP-2004
LOCUS BP224560 Sugano cDNA library, lymphocyte Daudi Homo sapiens cDNA
DEFINITION BP224560 Sugano cDNA library, lymphocyte Daudi Homo sapiens cDNA
clone DAT01067, mRNA sequence.
ACCESSION BP224560
VERSION BP224560.1 GI:52097465
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 581)
Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
FEATURES
source
1 .581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DAT01067"
/cell_type="lymphocyte"
/cell_lines="Daudi"
/clone_lib="Sugano cDNA library, lymphocyte Daudi"
/note="Burkitt's lymphoma"

ORIGIN
Query Match 35.9%; Score 579.4; DB 5; Length 581;
Best Local Similarity 99.8%; Pred. No. 1.2e-139;
Matches 580; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 548 TTAACATAGACAAATACACGGCATCGTCTGTGCGCGGAGATGGTATGTTACGCGAGG 607
Db 1 TTAACATAGACAAATACACGGCATCGTCTGTGCGCGGAGATGGTATGTTACGCGAGG 60
```

```
Qy 608 TGCTGCACCGTCTGATTGGGAGGACGCGAGGAGCGCGCGGTCGACCAGAACCAACCCCC 667
Db 61 TGCTGCACCGTCTGATTGGGAGGACGCGAGGAGCGCGCGGTCGACCAGAACCAACCCCC 120
Qy 668 GGGCTGTGTCGGTCCCAGTAGCCTCCGGATTGGAAATCATTTCCGCGAGGGTCAACGACT 727
Db 121 GGGCTGTGTCGGTCCCAGTAGCCTCCGGATTGGAAATCATTTCCGCGAGGGTCAACGACT 180
Qy 728 GCGTGTGTTACTCCACCGTGGGACACGCGAGCAAACTCGGCGCTGCATATCGTTG 787
Db 181 GCGTGTGTTACTCCACCGTGGGACACGCGAGCAAACTCGGCGCTGCATATCGTTG 240
Qy 788 TTGGGGACTCGCTGGCCATGATGTCTTCAGTCCACCAACAGCAGACACTCTCTCGCT 847
Db 241 TTGGGGACTCGCTGGCCATGATGTCTTCAGTCCACCAACAGCAGACACTCTCTCGCT 300
Qy 848 ACTCCGTGTCCTCTGCTGGGCTACGGCTTTACGGGGACATCATCAAGGACAGTGAGAAGA 907
Db 301 ACTCCGTGTCCTCTGCTGGGCTACGGCTTTACGGGGACATCATCAAGGACAGTGAGAAGA 360
Qy 908 AACGTTGGTGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGACCTTCTCTCCACC 967
Db 361 AACGTTGGTGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGACCTTCTCTCCACC 420
Qy 968 ACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACAAACACGCGTGGGATCTCCAAAGG 1027
Db 421 ACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACAAACACGCGTGGGATCTCCAAAGG 480
Qy 1028 ATAGGAAGCCCTGCGGGCAGGATGCTTTTTCAGGCAAAAGCAAGCAGCTGGAGG 1087
Db 481 ATAGGAAGCCCTGCGGGCAGGATGCTTTTTCAGGCAAAAGCAAGCAGCTGGAGG 540
Qy 1088 AGGACGAGAGAAAGCACTGTATGTTTGAAGCTGCGGAG 1128
Db 541 AGGACGAGAGAAAGCACTGTATGTTTGAAGCTGCGGAG 581

RESULT 14
BP310011 584 bp mRNA linear EST 17-SEP-2004
LOCUS BP310011 Sugano cDNA library, brain Homo sapiens cDNA clone
DEFINITION BP310011 Sugano cDNA library, brain Homo sapiens cDNA clone
NR001274, mRNA sequence.
ACCESSION BP310011
VERSION BP310011.1 GI:52238986
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 584)
Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
FEATURES
source
1 .584
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NR001274"
/tissue_type="brain"
/clone_lib="Sugano cDNA library, brain"

ORIGIN
Query Match 35.9%; Score 579.2; DB 5; Length 584;
Best Local Similarity 99.5%; Pred. No. 1.4e-139;
```



Matches 581; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
QY 415	CCGTTTGGAGGAAAAGGCAAGCAAGCGGATATATGAAGAAAGTGGCACCACCTGTTTC 474
Db 1	CCGTTTGGAGGAAAAGGCAAGCGGCGGATATATGAAGAAAGTGGCACCACCTGTTTC 60
QY 475	ACCTTTAGCCTCCATCACCACCTCACATCATCTGTTACTGAAACATGCTAAATCAGGCCCAAGGAG 534
Db 61	ACCTTTAGCCTCCATCACCCTGACATCATCTGTTACTGAAACATGCTAAATCAGGCCCAAGGAG 120
QY 535	ACTCTGTATGAGATTAAACATAGACAATAACGACCGCATCTGTCGTGTCGGCGGAGATGGT 594
Db 121	ACTCTGTATGAGATTAAACATAGACAATAACGACCGCATCTGTCGTGTCGGCGGAGATGGT 180
QY 595	ATGTTTCAGGAGGCTGTCACCGTCTGATTGGGAGGACGACAGAGAGCGCCGGGTCGAC 654
Db 181	ATGTTTCAGGAGGCTGTCACCGTCTGATTGGGAGGACGACAGAGAGCGCCGGGTCGAC 240
QY 655	CAGAACCAACCCCGGGCTGTGTCGTGCCAGTACGCTCCGGATTGGGAATCAATCCCGCA 714
Db 241	CAGAACCAACCCCGGGCTGTGTCGTGCCAGTACGCTCCGGATTGGGAATCAATCCCGCA 300
QY 715	GGGTCAACCGGACTGGCTGTGTTACTCCACCGTGGGCACCGACGACGAGAAACCTTCGGCG 774
Db 301	GGGTCAACCGGACTGGCTGTGTTACTCCACCGTGGGCACCGACGACGAGAAACCTTCGGCG 360
QY 775	TGCGATATCGTTGTTGGGAGCTCGCTGGCCATGAGATGTCTCATGTCACCAACACAGC 834
Db 361	CTGCATATCGTTGTTGGGAGCTCGCTGGCCATGAGATGTCTCATGTCACCAACACAGC 420
QY 835	ACACTCCCTTCGCTACTCCGTCGTCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAG 894
Db 421	ACACTCCCTTCGCTACTCCGTCGTCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAG 480
QY 895	GACAGTGAGAAAGAACGGTGGTGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGACC 954
Db 481	GACAGTGAGAAAGAACGGTGGTGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGACC 540
QY 955	TTCTCTCTCCACCATCTGCTATGAAGGACAGTGTCTTCTCTCCCT 998
Db 541	TTCTCTCTCCACCATCTGCTATGAAGGACAGTGTCTTCTCTCCCT 584

RESULT 15  
BP309990  
LOCUS BP309990 Sugano cDNA library, brain Homo sapiens cDNA clone  
DEFINITION NRR01216, mRNA sequence.  
ACCESSION BP309990  
VERSION BP309990.1 GI:52238965  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 584)  
Mizushima-Sugano J., Nakai K. and Sugano S.  
Suzuki Y., Yamashita R., Shiota M., Sakakibara Y., Chiba J.,  
Sequence comparison of human and mouse genes reveals a homologous  
block structure in the promoter regions  
Genome Res. 14 (9), 1711-1718 (2004)  
Contact: Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
Email: ysuzuki@ims.u-tokyo.ac.jp.

FEATURES  
source  
1..584  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="NRR01216"  
/tissue\_type="brain"

/clone\_lib="Sugano cDNA library, brain"

ORIGIN	
Query Match 35.3%; Score 570.4; DB 5; Length 584; Best Local Similarity 99.5%; Pred. No. 2.7e-137; Matches 582; Conservative 0; Mismatches 2; Indels 1; Gaps 1;	
QY 415	CCGTTTGGAGGAAAAGGCAAGCGGATATATGAAGAAAGTGGCACCACCTGTTTC 474
Db 1	CCGTTTGGAGGAAAAGGCAAGCGGCGGATATATGAAGAAAGTGGCACCACCTGTTTC 59
QY 475	ACCTTTAGCCTCCATCACCACCTGACATCATCTGTTACTGAAACATGCTAAATCAGGCCCAAGGAG 534
Db 60	ACCTTTAGCCTCCATCACCACCTGACATCATCTTACTGAACATGCTAAATCAGGCCCAAGGAG 119
QY 535	ACTCTGTATGAGATTAAACATAGACAATAACGACCGCATCTGTCGTGTCGGCGGAGATGGT 594
Db 120	ACTCTGTATGAGATTAAACATAGACAATAACGACCGCATCTGTCGTGTCGGCGGAGATGGT 179
QY 595	ATGTTTCAGGAGGCTGTCACCGTCTGATTGGGAGGACGACAGAGAGCGCCGGGTCGAC 654
Db 180	ATGTTTCAGGAGGCTGTCACCGTCTGATTGGGAGGACGACAGAGAGCGCCGGGTCGAC 239
QY 655	CAGAACCAACCCCGGGCTGTGTCGTCCCAGTAGCCTCCGGATTGGGAATCAATCCCGCA 714
Db 240	CAGAACCAACCCCGGGCTGTGTCGTCCCAGTAGCCTCCGGATTGGGAATCAATCCCGCA 299
QY 715	GGGTCAACCGGACTCGCTGTGTTACTTCCACCGTGGGCACCGACGACGAGAAACCTTCGGCG 774
Db 300	GGGTCAACCGGACTCGCTGTGTTACTTCCACCGTGGGCACCGACGACGAGAAACCTTCGGCG 359
QY 775	CTGCATATCGTTGTTGGGACTCGCTGGCCATGATGTGTCCTCAGTCCACCAACACAGC 834
Db 360	CTGCATATCGTTGTTGGGACTCGCTGGCCATGATGTGTCCTCAGTCCACCAACACAGC 419
QY 835	ACACTCCCTTCGCTACTCCGTCGTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAG 894
Db 420	ACACTCCCTTCGCTACTCCGTCGTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAG 479
QY 895	GACAGTGAGAAAGAACGGTGGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACC 954
Db 480	GACAGTGAGAAAGAACGGTGGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACC 539
QY 955	TTCTCTCTCCACCATCTGCTATGAAGGACAGTGTCTTCTCTCCCT 999
Db 540	TTCTCTCTCCACCATCTGCTATGAAGGACAGTGTCTTCTCTCCCT 584

Search completed: September 6, 2005, 02:19:47  
Job time : 5907.4 secs

***This Page Blank (uspto)***

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2005, 19:28:44 ; Search time 184.39 Seconds  
(without alignments)  
2892.921 Million cell updates/sec

Title: US-10-631-958-2

Perfect score: 1717

Sequence: 1 PKHLLVFNPFGKCGQKRI.....KCSRPNFLRLHRTNQDQ 326

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10631958/runat 02092005.165811.4432/app query.fasta\_1.1941  
-DB=Issued Patents NA -QFMT=fastap -SUPFLX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=ptc -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAFSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10631958 @CGN 1.1.397 @runat 02092005.165811.4432 -NCFU=6 -ICFU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1633.5	95.1	4432	4	US-09-774-528-148
2	613	35.7	2064	4	Sequence 148, App
3	566	33.0	1084	4	Sequence 14306, A
4	243	14.2	1857	4	Sequence 15155, A
5	243	14.2	2280	4	Sequence 3, Appli
6	237.5	13.8	2598	4	Sequence 13, Appl
7	228	13.3	1783	4	Sequence 11, Appl
8	227.5	13.2	1155	4	Sequence 1155, Ap
9	227.5	13.2	1205	4	Sequence 1, Appli
10	226	13.2	1149	4	Sequence 1, Appli
11	220.5	12.8	1533	4	Sequence 5, Appli
12	216	12.6	1611	4	Sequence 90, Appl
					Sequence 1756, Ap

13	181.5	10.6	1875	4	US-09-614-221A-399
14	173.5	10.1	2462	4	Sequence 399, App
15	172	10.0	901	4	Sequence 796, App
16	171.5	10.0	1764	4	Sequence 30448, A
17	135.5	7.9	1012	4	Sequence 3940, Ap
18	134.5	7.8	1239	4	Sequence 12677, A
19	132	7.7	942	4	Sequence 263, App
20	125	7.3	485	4	Sequence 3500, Ap
21	125	7.3	885	4	Sequence 31476, A
22	125	7.3	948	4	Sequence 1079, Ap
23	118.5	6.9	3089	4	Sequence 518, App
24	117	6.8	915	4	Sequence 188, App
25	116.5	6.8	900	3	Sequence 1961, Ap
26	110	6.4	912	4	Sequence 332, App
27	110	6.4	4411529	3	Sequence 1479, Ap
28	109	6.3	498	4	Sequence 1, Appli
29	109	6.3	978	4	Sequence 59, Appl
30	106.5	6.2	7195	4	Sequence 3132, Ap
31	106.5	6.2	7198	4	Sequence 12897, A
32	106	6.2	951	4	Sequence 15682, A
33	106	6.2	975	3	Sequence 1349, Ap
34	106	6.2	3761	4	Sequence 1937, Ap
35	105	6.1	13542	4	Sequence 3656, Ap
36	105	6.1	13542	4	Sequence 408, App
37	105	6.1	13542	4	Sequence 154, App
38	101	5.9	1393	4	Sequence 154, App
39	100.5	5.9	4403765	3	Sequence 40, Appli
40	100	5.8	1003	3	Sequence 2, Appli
41	98	5.7	3342	4	Sequence 1064, Ap
42	96	5.6	2085	4	Sequence 748, App
43	95.5	5.6	5381	3	Sequence 1002, Ap
44	95.5	5.6	81001	3	Sequence 4, Appli
45	95	5.5	191	4	Sequence 1, Appli
					Sequence 28485, A

ALIGNMENTS

RESULT 1  
US-09-774-528-148  
; Sequence 148, Application US/09774528  
; Patent No. 6743619  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Zhou, Ping  
; APPLICANT: Goodrich, Ryle  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Zhang, Jie  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6743619el Nucleic Acids and  
; FILE REFERENCE: 802  
; CURRENT APPLICATION NUMBER: US/09/774,528  
; CURRENT FILING DATE: 2001-01-30  
; NUMBER OF SEQ ID NOS: 441  
; SOFTWARE: pt\_Fl\_genes Version 2.0  
; SEQ ID NO 148  
; LENGTH: 4432  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1129) .. (2817)  
US-09-774-528-148  
Alignment Scores:

```
Pred. No.: 8.62e-208 Length: 4432
Score: 1633.50 Matches: 313
Percent Similarity: 96.32% Conservative: 1
Best Local Similarity: 96.01% Mismatches: 11
Query Match: 95.14% Indels: 11
DB: 4 Gaps: 1

US-10-631-958-2 (1-326) x US-09-774-528-148 (1-4432)

Qy 1 ProlysHisLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
Db 1591 CCAAAGCATTTACTGGTATTTATCAACCGTTTGGAGAAAAGCAAGCAAGCGGATA 1650

Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
Db 1651 TATGAAAGAAAGTGGCAACACTGTTCCACTTACCTCCATCACCACCTGACATCATC--- 1707

Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 1708 -----GTTACTGAACATGCTAATCAGGCCCAAGGAG 1737

Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyAspGly 80
Db 1738 ACTCTGTATGATTAACATAGACAAATACACGGCATCGTCTGTGTCGGCGAGATGGT 1797

Qy 81 MetPheSerGluValLeuHisGlyIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 1798 ATGTTACGAGAGTGTCTGACGGTCTGATTTGGAGGACGACAGAGCGCGGGTGCAC 1857

Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAla 120
Db 1858 CAGAACACCCCGGGCTGTGTCGCCAGTAGCTCCGATTCGGATTTGGAATCATTCGCCGA 1917

Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 1918 GGGTCAACGGACTCGTGTGTATCTCCACCGTGGGACACGACGACGAGAAACCTCGCG 1977

Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
Db 1978 CTGCATATCGTTTGGGACTCGTGGCCATGATGTCTCTCAGTCCACCAACACAGC 2037

Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db 2038 ACACCTCTCTGCTACTCGTGTCTGCTGGCTACGGCTTCTACGGGGACATCATCAAG 2097

Qy 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 2098 GACAGTGAGAGAAACGGTGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAGACC 2157

Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 2158 TTCCTCTCCACCACTGCTATGGAGGACAGTGTCTCTCTCTCCCTGCACACACACCGGTG 2217

Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db 2218 GGATCTCAAGGATAGGAAGCCCTGCGCGGACAGTGTCTTGTGTTGCGAGCAAGCAAG 2277

Qy 241 GlnGlnLeuGluGluGlnLysIleValLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 2278 CAGCAGCTGGAGGAGGACGAGAAAGCATGTATGTTTGGAGCTGCGGAGGACGTG 2337

Qy 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 2338 GAGGAGTGGCAAGTCTCTGTGGGAGTTTCTGGCCATCAATGCCACAAACATGTCCTGT 2397

Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
Db 2398 GCTTGTCCCGGAGGCCCGCGGCTCTCCCGGCTGCGCCACTTGGGAGAGCGGTCTTCT 2457

Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
Db 2458 GACCTCTCTCATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAC 2517
```

```
Qy 321 ThrAsnGlnGlnAspGln 326
Db 2518 ACCAACCCAGCAGGACCAG 2535

RESULT 2
US-09-270-767-14306
; Sequence 14306, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14306
; LENGTH: 2064
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-14306

Alignment Scores: 3.06e-71 Length: 2064
Pred. No.: 613.00 Matches: 134
Percent Similarity: 50.66% Conservative: 59
Best Local Similarity: 35.17% Mismatches: 108
Query Match: 35.70% Indels: 80
DB: 4 Gaps: 7

US-10-631-958-2 (1-326) x US-09-270-767-14306 (1-2064)

Qy 2 LysHisIleLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyr 21
Db 610 CGCGCCCTTTTGGTCTTTATAACCCCTATGGAGTTCGCAAGCGGAGCTCAGACCTAT 669

Qy 22 GluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGlyAsn 41
Db 670 GAGCGCCATGTGAGACCTATTTCCAGCTTTCGCGGCTAGACGCCACGTGTATC----- 723

Qy 42 LysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGluThr 61
Db 724 -----ACCACCTCAGAGGGCAACCAAGTGAAGGACATA 756

Qy 62 LeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyWet 81
Db 757 CTTCTGAGCCATGATCTGGAGATATACGATCGGTTTGTCTGTCGGAGCGATGGCACC 816

Qy 82 PheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGln 101
Db 817 GTAGCAGAGGTATCAACGGACTGATATTCGTCMAATCGAGAGTTGGGACTGGACGAA 876

Qy 102 AsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGly 121
Db 877 CAGCGGCCACCATACATTCCAGAGCCG---GCTCTGCCAGTGGGTGTGATTCGCGCTGGC 933

Qy 122 SerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeu 141
Db 934 AGCACCCAGACCATTCGTTATAGTATGCACGGCAGCGGATGTGAGGACAGCGGCTATC 993

Qy 142 HisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThr 161
Db 994 CATGTGATTCCTGGGCCAGCATCGGGATTGGATGTGTGTCAGTGTGAGCAATGGGCACTCC 1053

Qy 162 LeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAsp 181
Db 1054 CTGCTCAGATCTCTGCGCAGTCTCTGAGTACGGGTACCTGGCGGATGTGGCAGCCAG 1113

Qy 182 SerGluLysLysArgTyrPheLeuAlaArgTyrAspPheSerGlyLeuLysThrPhe 201
Db 1114 AGCGAGAACTACCGCTGGATGGGACCGCGCGGTACAGTACAGTGGCGGTCAAGCCCTC 1173

Qy 202 LeuSerHisHisCysTyrGluGlyThrValSerPheLeu----- 214
```

978	GAGCGCATGTGAGACCTATTTTCCAGCTTCCGGCGGTAGAGCCACCTGTATC-----	925
Qy	LysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGluThr	61
Db	-----ACCACCTCAGAGGGCAACCAAGTAGTGAGGACATA	892
Qy	LeuTyrGluIleAsnIleAspLysTyrAspClyIleValCysValGlyIleAspClyMet	81
Db	CTTCTGAGCCATGATCTGGGAGTATACGATCGGTTGCTGTGCGGAGCGATGGCAC	832
Qy	PheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGln	101
Db	GTAGCAGAGTTCATCAACGGACTGATATTCGTCAAATGCCAGAGTTCGGACTGCACGA	772
Qy	AsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGly	121
Db	CAGCGGCGCACCATATCCAAAGACGG---GCTCGCCAGTGGGTGTGATTTCCCGCTGGC	715
Qy	SerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeu	141
Db	AGCACCGACCACTTGGTATAGTATGCACGGCAGCGCGGATGTGAGCAGCGGCTATC	655
Qy	HisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisAsnSerThr	161
Db	CATGTGATTTCTGGGCCAGCATCGGGGATTGGATGTGTGCAGTGTGAGCAATGGCCAGTCC	595
Qy	LeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAsp	181
Db	CTGCTCAGATCTGTGCCAGTGTCTGAGCTACCGGTACCTGGCGGATGTGGCAGCCGAC	535
Qy	SerGluLysLysArgTyrLeuAlaArgTyrAspPheSerGlyLeuLysThrPhe	201
Db	AGCGAGAACTACCGCTGGATGGGACCGCGCGGTACGAGTACAGTGGGCGTCAAGGCCCTTC	475
Qy	LeuSerHisHisCysTyrGluGlyThrValSerPheLeu-----	214
Db	CTGAATAATCGCGGCTATGACCGCGGAACGTAGAATGTTAGAGAGAGCCCGCATCTTCTACTGT	415

## C77 - - - -

414	ACACAGCGCGTGGAGGACATTCCGCAGAGTCCGATAGTGTGCTCGCTCGGAGAGTCA	355
Qy	----	226
Db	GTGCCATCCGTCTGCTATGCAATTGCCAGCGCTGCAGCTTCGCCAGCAGCATACAGGAA	295
Qy	ArgGlnSerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAla	256
Db	CAGCGATCCTCATTTGTTTCATCCAGAGGAAATCTAAAGAGGCA-----GAGCGC	247
Qy	AlaGluAspValGlu-----	261
Db	AATCAGCAGGTAGAAACAGAGGACTCTCATCTAGCGCCAGTGAAGCAGCACTCCTTAGG	187
Qy	-----	261

[illegible]

Qy 298 Gly 298  
|||  
Db 6 GGT 4

RESULT 4  
US-09-970-516-3  
; Sequence 3, Application US/09970516

```
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1857
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1857)
; OTHER INFORMATION:
US-09-970-516-3
Alignment Scores:
Pred. No.: 1,34e-21 Length: 1857
Score: 243.00 Matches: 67
Percent Similarity: 47.41% Conservative: 43
Best Local Similarity: 28.88% Mismatches: 90
Query Match: 14.15% Indels: 32
DB: 4 Gaps: 4

US-10-631-958-2 (1-326) x US-09-970-516-3 (1-1857)
Qy 1 ProLysHisLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgile 20
Db 430 CGGCCCCGGTGTCTTATGGTCAATCCCTTTGGGGGTGGGGCTGGCCCTGGCAGTGG 489
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleGly 40
Db 490 TGTAAAGAACCAAGCTGCTCCCATGATCTCTGAAGCTGGGCTGTCTTCAACCTCATCCAG 549
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 550 -----ACAGAACGACAGAACCCGCGGGAG 576
Qy 61 ThrLeuTyrGluIleAsnLysTyrAspGlyIleValCysValGlyLysArgile 80
Db 577 CTGTGTCAGGGGCTGAGCTGAGTGAAGTGGATGCGTCACGGTCTCGGAGACGGG 636
Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValasp 100
Db 637 CTGCTCCATGAGGTGCTGAACGGGCTCTAGATCGC----- 672
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAla 120
Db 673 CTGACTGGGAGGAGCTGTGAAGATGCT-----GTGGGCACTCTCCCTGCG 720
Qy 121 GlySerThrAspCysValCysTyrSer-----ThrVal 131
Db 721 GGCTCGGCAAGCGCTGCGGAGCAGTGAACAGCAGCGGGGATTTGAGCCAGCCCTG 780
Qy 132 GlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMet 151
Db 781 GGCTCGACCTGTTGCTCAACTGCTCACTGTGCTGTCGGGGTGGTGGCCACCCACTG 840
Qy 152 AspValSerSerValHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGly 171
Db 841 GACCTGCTCTCGTGACCGCTGCGGCTCCCGCTGTTCTCTCTCTCTCTGTGTGGCC 900
Qy 172 TyrGlyPheTyrGlyAspIleIleLysAspSerGluLysLysArgTyrTrpLeuGlyLeuAla 191
Db 901 TGGGGCTTCGTGTAGATGTGATATCCAGAGCAGCGCTTCAGGGCTTGGGAGTGGC 960
Qy 192 ArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisCysTyrGluGlyThrVal 211
Db 961 CGCTTACACTGGGACGGTGTGTGGGCTCGCCACACTGCACCTACCGCGGAGCGCTC 1020
```

```
Qy 212 SerPheLeuProAlaGlnHisThrValGlySerPro 223
Db 1021 TCCTACCTCCCGCCCACTGTGGAACCTGCTCGCC 1056
```

## RESULT 5

```
US-09-817-676A-13
; Sequence 13, Application US/09817676A
; Patent No. 6800470
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817,676A
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/194,318
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 2380
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1860)
; PUBLICATION INFORMATION:
; TITLE: Molecular cloning and functional characterization of a
; JOURNAL: J. Biol. Chem.
; VOLUME: 275
; ISSUE: 26
; PAGES: 19513-19520
; DATE: 2000-06-30
; DATABASE ACCESSION NUMBER: AF245447
; DATABASE ENTRY DATE: 2000-06-27
US-09-817-676A-13
```

```
Alignment Scores:
Pred. No.: 2,09e-21 Length: 2380
Score: 243.00 Matches: 67
Percent Similarity: 47.41% Conservative: 43
Best Local Similarity: 28.88% Mismatches: 90
Query Match: 14.15% Indels: 32
DB: 4 Gaps: 4
```

US-10-631-958-2 (1-326) x US-09-817-676A-13 (1-2380)

```
Qy 1 ProLysHisLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgile 20
Db 436 CGGCCCCGGTGTCTTATGGTCAATCCCTTTGGGGGTGGGGCTGGCCCTGGCAGTGG 495
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleGly 40
Db 496 TGTAAAGAACCAAGCTGCTCCCATGATCTCTGAAGCTGGGCTGTCTTCAACCTCATCCAG 555
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 556 -----ACAGAACGACAGAACCCGCGGGAG 582
Qy 61 ThrLeuTyrGluIleAsnLysTyrAspGlyIleValCysValGlyLysArgile 80
Db 583 CTGTGTCAGGGGCTGAGCCCTGAGTGGATGGCATCGTCACGGTCTCGGAGACGGG 642
Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValasp 100
Db 643 CTGCTCCATGAGGTGCTGAACGGGCTCTAGATCGC----- 678
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAla 120
Db 679 CCGTACTGGGAGGAGCTGTGAAGATGCT-----GTGGGCACTCTCCCTGCG 726
```

Qy	24	LysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGlyAsnLysPhe	43
		: : :     : : : : :	
Db	888	CACGTGGTCCAAATGATCTCTGAAGCTGGSGCTCCTTCAACCTCATACAG-----	938
Qy	44	TyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGluThrLeuTyr	63
		: : :     : : : : :	
Db	939	-----ACAGAACGACAGAACCATGCCCGTAGCTGGTGTCGAC	974
Qy	64	GluIleAsnIleAspLysTyrAspGlyIleValCysvalGlyAspGlyMetPheSer	83
		: : : : : : : : : : : : : : : : :	
Db	975	GGGTTAAGCCCTAGTGAAGGAGCATTTGCATCTGTCTGGACGCGGCTCTTAC	1034
Qy	84	GluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHis	103
		: : : : :     : : : : :	
Db	1035	GAGTGTCTGAATGGGCTCTTGATCGG-----CCAGACTGG	1070
Qy	104	ProArgalaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySerThr	123
		: : :     : : : : :	
Db	1071	GAGATGCCGTGGGATGCC-----ATTGGTGTCTCCCTCTGATCGGC	1118
Qy	124	AspCysValCys-----TyrSerThrValGlyThrSer	134
		: : : : : : : : : : : : : : : : :	
Db	1119	AATCGCTACTCGGGCGGTGAGCCATCATGGCGGGTTTTGACGACAGTGTGTGGTGTGAC	1178

**RESULT 6**

[illegible]

```

? CORRENT FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FastSEQ for Windows Version 4.0
? SEQ ID NO 1155
? LENGTH: 1783
? TYPE: DNA
? ORGANISM: Human
US-09-949-016-1155

```





Db 451 AACCTGCTCTCTGCACACGGCTTCGGGGCTGGCGCTCTTCTCTGTGCTCAGCCTGGCC 510  
Qy 172 TyrGlyPheTyrGlyAspIleIleLysAspSerGlyLysLysArgTrpLeuGlyLeuAla 191  
Db 511 TGGGGCTTCATTGCTGATGTGGACCTAGAGAGTGAAGAATATCGCGCTCTGGGGGAGATG 570  
Qy 192 ArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrVal 211  
Db 571 CGCTTCACTCTGGGCACCTTCTGCTGCTGGCAGCGCTCGGCACCTACCGCGCGCGACTG 630  
Qy 212 SerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysPro 228  
Db 631 GCCTACCTCCCTGTAGGAAGA---GTGGGTTCCAAGACACCTGCCTCCCCC 678

## RESULT 9

US-09-959-897-1  
; Sequence 1, Application US/09959897  
; Patent No. 6730480  
; GENERAL INFORMATION:  
; APPLICANT: PITSON, Stuart M  
; APPLICANT: Brian, WATTENBERG W  
; APPLICANT: Pu, XIA  
; APPLICANT: Richard, D'ANDREA J  
; APPLICANT: Jennifer, BABLE R  
; APPLICANT: Mathew, VADAS A  
; TITLE OF INVENTION: SPHINGOSINE KINASE ENZYME  
; FILE REFERENCE: PITSON-1  
; CURRENT APPLICATION NUMBER: US/09/959,897  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: PCT/AU00/00457  
; PRIOR FILING DATE: 2000-05-12  
; PRIOR APPLICATION NUMBER: AU PQ 0339  
; PRIOR FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: AU PQ 1504  
; PRIOR FILING DATE: 1998-07-08  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1205  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (33)..(1184)  
; OTHER INFORMATION:  
US-09-959-897-1

Alignment Scores:  
Pred. No.: 7,48e-20 Length: 1205  
Score: 227.50 Matches: 64  
Percent Similarity: 45.15% Conservative: 43  
Best Local Similarity: 27.00% Mismatches: 97  
Query Match: 13.25% Indels: 33  
DB: 4 Gaps: 5

US-10-631-958-2 (1-326) x US-09-959-897-1 (1-1205)

Qy 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyGlyGlnGlyLysArgIle 20  
Db 72 CCCTGCCGCTGCTGGTGTCTGAACCGCGCGGGCAAGGCAAGCCCTTGAGCTC 131  
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleGly 40  
Db 132 TTCCGGAGTCAGTGCAGGCCCTTTTGGCTGAGGCTGAATC----- 173  
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60  
Db 174 -----TCCTTCACGTGATGCTCACTAGCGCGCGCAACACCGCGGGAG 218  
Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80  
Db 219 CTGGTGGGTCGGAGAGCTGGGGCGCTGGGACGCTCTGGTGGTCACTGTGGAGACGGG 278

Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100  
Db 279 CTGATGCACGAGGTGGTGAACGGCTCATGAGCGCGCTGACTGGGAGACCGCCATCCAG 338  
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAla 120  
Db 339 AAG-----CCCTGTGTAGCTCCACGCA 362  
Qy 121 GlySerThrAspCysValCys-----TyrSerThrValGly 132  
Db 363 GGCTCTGCGCAACGGCTGGCAGCTTCTTGAACCATTTATGCTGGCTATGAGCAGGTACCC 422  
Qy 133 ThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAla---Met 151  
Db 423 AATGAAGACCTCTCGACCAACTGCACGCTATGTCTGCCGCGGCTGCTGTACCCCATG 482  
Qy 152 AspValSerSerValHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGly 171  
Db 483 AACCTGCTGCTCTGCACACGGCTTCGGGGCTGGCGCTTCTCTGTCTGCTCAGCCTGGCC 542  
Qy 172 TyrGlyPheTyrGlyAspIleIleLysAspSerGlyLysLysArgTrpLeuGlyLeuAla 191  
Db 543 TGGGGCTTCAATTGCTGATGTGCACCTAGAGAGTGAGAAGTATCGCGCTCTGGGGGAGATG 602  
Qy 192 ArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrVal 211  
Db 603 CGCTTCACTCTGGGCACCTTCTCGCTGCTGGCAGCTTTCGGCAGCTTACCGCGCGGACTG 662  
Qy 212 SerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysPro 228  
Db 663 GCTTACCTCCCTGTAGGAAGA---GTGGGTTCCAAGACACCTGCCTCCCCC 710

## RESULT 10

US-09-970-516-5  
; Sequence 5, Application US/09970516  
; Patent No. 6610534  
; GENERAL INFORMATION:  
; APPLICANT: No. 6610534artis AG  
; TITLE OF INVENTION: Induction of blood vessel formation through administration of  
; FILE REFERENCE: polynucleotides encoding sphingosine kinases  
; FILE REFERENCE: 4-31617  
; CURRENT APPLICATION NUMBER: US/09/970,516  
; CURRENT FILING DATE: 2001-10-04  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 1149  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1149)  
; OTHER INFORMATION:  
US-09-970-516-5

Alignment Scores:  
Pred. No.: 1.09e-19 Length: 1149  
Score: 226.00 Matches: 63  
Percent Similarity: 45.15% Conservative: 44  
Best Local Similarity: 26.58% Mismatches: 94  
Query Match: 13.16% Indels: 36  
DB: 4 Gaps: 5

US-10-631-958-2 (1-326) x US-09-970-516-5 (1-1149)

Qy 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyGlyGlnGlyLysArgIle 20  
Db 40 CCATGACAGGTGGTGGTCTGCTGAACCCCGAGGGTGGCAAGGCTCTGAGCTC 99  
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleGly 40  
Db 100 TTCCAGACCGCTGTGACGCCCTTCTTGGAGGAGGAGAGATAACCTTTAAACTGATA--- 156

```
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAenGlnAlaLysGlu 60
Db 157 -----CTCACCGAACGGAAGAACCATGCCAGGGAG 186
Qy 61 ThrLeuTyrGluIleAsnLysTyrAspGlyIleValCysValGlyGlyAspGly 80
Db 187 CTGGTGTGTGCAGAGGAGTTGGGTCACTGGGACGCCCTGGCAGTCATGTCGGGTGATGGT 246
Qy 81 MetPheSerGluValLeuHisGlyLeuGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 247 CTGATGATGAGGTGGTGAATGGGCTAATGAACGCCGCTGGGAGACGTCGCATCCAG 306
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db 307 AAA-----CCCTGTGTGAGCTCCCTGGA 330
Qy 121 GlySerThrAspCysValCys-----TyrSerThrValGly 132
Db 331 GGCTCCGCAATGGCTGGCAGCTTCTGTGAACCACTATGCTGGGTACGAGCAGGTGACT 390
Qy 133 ThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAla--Met 151
Db 391 AATGAAGACCTGCTCATCATGCACACTGCCTGTGTGCCCGCGCTGTACCCATG 450
Qy 152 AspValSerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGly 171
Db 451 AACCTGTCTCCCTGCACACTGCTTCTGGGCTGGCTCTATTCTGTGCTCAGTCTGTC 510
Qy 172 TyrGlyPheTyrGlyAspIleIleLysAspSerGluLysArgTyrTrpLeuGlyLeuAla 191
Db 511 TGGGGCTTTGTGTGCTGAGTGCAGTCTCAGAGTGAAGTACAGCGCGCTTGGGGGAGATT 570
Qy 192 ArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisCysTyrGluGlyThrVal 211
Db 571 CGTTTCACATGGGCACCTCTTTCGCTCAGCAAGCTGCCCATCTACCAAGGCCAATG 630
Qy 212 SerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysPro 228
Db 631 GCCTACCTTCCT-----GTAGGAAGTGGGCTCTAAGAGACCC 669

RESULT 11
US-09-205-258-90
; Sequence 90, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007F1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 90
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (12)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (123)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
```

```
; LOCATION: (1522)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1527)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-205-258-90

Alignment Scores:
Pred. No.: 1e-18 Length: 1533
Score: 220.50 Matches: 70
Percent Similarity: 45.80% Conservative: 39
Best Local Similarity: 29.41% Mismatches: 95
Query Match: 12.84% Indels: 34
DB: 4 Gaps: 7

US-10-631-958-2 (1-326) x US-09-205-258-90 (1-1533)

Qy 1 ProLysHisLeuLeuValPheIleAenProPheGlyGlyGlyGlnGlyLysArgile 20
Db ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
112 CCTGCGCGTGNCTGGTGTCTGAACCCGCGCGCGCAAGGGCAAGGCTTGAGCTC 171
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
172 TTCGGAGTCAGTCAGCCCTTTGGCTGAGGCTGAATC----- 213
Qy 41 AenLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
214 -----TCCTTCACGCTGATGCTCACTGAGCGCGGAAACACGCGGGGAR 258
Qy 61 ThrLeuTyrGluIleAenLysPheTyrAspGlyIleValCysValGlyGlyAspGly 80
Db ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
259 CTGGTGGCGTGGAGGAGTCGGCGCTGGRACGCTCTGGTGTGATGTGGAGACGGG 318
Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
319 CTGATGCAGAGTGTGACGGGCTTCATGGA-----GCCGCTGACGGGA 366
Qy 101 GlnAenHisProArgAlaValLeuValPro-SerSerLeuArgIleGlyIleIleProAl 120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
367 GACCGCCATCCAGAGCC-----CCTGTGTAGCTC-----CCAGC 402
Qy 120 aGlySerThrAspCysValCys-----TyrSerThrValG1 132
Db ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
403 AGGCTCTGGCAACGCSCTGGCAGCTTCCTTAAACCAATTATGCTGGCTATRAGCAGSTCAC 462
Qy 132 yThrSerAspAlaGluThrSerAlaLeuHisIleValGlyAspSerLeuAla--Me 151
Db ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||
463 CAATGAAGACCTCTGACCACTGACCGTATGCTGTGCCCGCGCTGTGTACCCAT 522
Qy 151 tAspValSerSerValHisHisAenSerThrLeuLeuArgTyrSerValSerLeuLeuG1 171
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
523 GAACCTGTGTCTCTGCACACGGCTTCGGGGTGGCGCTCTTCTCTGTGCTCAGCCTGCG 582
Qy 171 yTyrGlyPheTyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAl 191
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
583 CTGGGGCTTCATTGTGTGATGTGGACCTAGAGAGTGAGAAGTATCGCGCTCTGGGGGAGAT 642
Qy 191 aArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrVa 211
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
643 GCGCTTCACTCTGGGACCTCTCTCGCTGGCAGCCCTGGCAGCTACCGCGCGCGACT 702
Qy 211 lSerPheLeuProAlaGlnHisThrValGlySerProArgAspAGlyLysPro 228
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
703 GGCCTACCTCTGTAGGAGA--GTGGTTCCAGACACCTGCTCCCCC 751

RESULT 12
US-09-248-796A-1756
; Sequence 1756, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
```

```
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248.796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 1756
; LENGTH: 1611
; TYPE: DNA
; ORGANISM: Candida albicans
US-03-248-796A-1756

Alignment Scores:
Pred. No.: 4.41e-18 Length: 1611
Score: 216.00 Matches: 79
Percent Similarity: 42.95% Conservative: 58
Best Local Similarity: 24.76% Mismatches: 131
Query Match: 12.58% Indels: 52
DB: 4 Gaps: 10

US-10-631-958-2 (1-326) x US-09-248-796A-1756 (1-1611)

Qy 4 LeuLeuValPheIleAenProPheGlyGlyGlyGlnGlyLysArgIleTyrGluArg 23
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
493 ATTTTGGTATTGATAAACCCGCGCATGGCGGCCAAGCACACGCAAACTATTACAAAAAT 552
Qy 24 LysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGlyAsnLysPhe 43
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
553 AAAATCTTACCAATATTACAAAGCGCT-----CGTGCT 585
Qy 44 TyrValAenTyrValGluValIleThrGluHisAlaAenGlnAlaLysGluThrLeuTyr 63
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
586 AATGTACGATTATTGAA-----ACTAAATATCATGGACACGCCACTGAGATTCCGCGT 639
Qy 64 GluLeuAenIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPheSer 83
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
640 GAGCTAGATGTCATGATTATGATAATAATTGTTGTTGTTCTGCGCATGGGATACCTCAT 699
Qy 84 GluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHis 103
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
700 GAAGTATCAATGGCTCTCTATCTTCCACAGAT----- 732
Qy 104 ProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySerThr 123
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
733 ---AAAGGTTTACTGGCATTCACAAAATTCAGTTACTCAATTACCTTGTGGGTACGGG 789
Qy 124 AspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIle 143
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
790 AACGCGTTGAGTTTGAGTACACATGTTAGTAAAAATGCTTCAGTTGCAACTCTTTATATG 849
Qy 144 ValValGlyAspSerLeuAlaMetAspValSerSerValHisHis-----AsnSer 160
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
850 TTGAAGCTCATAGACAAAATTCGATTATGATGCTATTACCCAAAGGTACAGGAAGTAA 909
Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleLys 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
910 AAGATAAGAAATTTGTCAATTTTAAAGTCAGTGTACGGTATATTGCTGATTCTGATATT 969
Qy 181 AspSerGluLysLysArgTrpLeuAlaArgTyrAsp----- 194
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
970 GGAACAGAACATTTACGTTGGTGGTCCCTATACGGTTTGAACCTGGAGTGATACAAAAA 1029
Qy 195 ---PheSerGlyLeuLys-----ThrPheLeuSerHisHisCysTyrGluGly 209
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1030 GTGTCTCGGGGCAAAATATCTCTGTGACCTATTTGTGAAATACAAAG---TACGATAAT 1086
Qy 210 ThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysProCys 229
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1087 AATTGAGAGATTTTGAATCATGTAAATGATTACTTGTAGTAATATGAT----- 1134
```

Qy 230 ArgAlaGlyCysPheValCysArgGlnSerIysGlnGlnLeuGluGluGlnLysLys 249  
Db 1135 -----ACTGAAACAGTACCCTTGTCTACCTGAAGACAACTTG 1173  
Qy 250 AlaLeuTyrGlyLeuGluAlaAlaGluaspVal---GluGluTTPGln-ValValCysG1 268  
Db 1174 CAATATACGACCCCTGATTAGATCAACCTGTCCCAATGATTGGAGACATATTCCTCAA 1233  
Qy 268 YLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArg-- 287  
Db 1234 GAAATTTCTCATATTTGAAACATCTTGTATGTTGGCAAAATGCCATTCGTCTCGAGAT 1293  
Qy 288 ----GlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuLeu 304  
Db 1294 ACTCAGTTTTTC-CCTGCTGCTCTACCGAACGATGGTTCAATGACATGATTGTC 1347

## RESULT 13

US-09-614-221A-399  
; Sequence 399, Application US/09614221A  
; Patent No. 6723837  
; GENERAL INFORMATION:  
; APPLICANT: Karunanandaa, Balasulojini  
; APPLICANT: Yu, Jaehyuk  
; APPLICANT: Kishore, Ganesh M.  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED  
; TITLE OF INVENTION: WITH STEROL SYNTHESIS AND METABOLISM  
; FILE REFERENCE: 16516.075  
; CURRENT APPLICATION NUMBER: US/09/614,221A  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/142,981  
; PRIOR FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 626  
; SEQ ID NO 399  
; LENGTH: 1875  
; TYPE: DNA  
; ORGANISM: Saccharomyces cerevisiae

## US-09-614-221A-399

Alignment Scores:  
Pred. No.: 2,51e-13 Length: 1875  
Score: 181.50 Matches: 62  
Percent Similarity: 42.68% Conservative: 43  
Best Local Similarity: 25.20% Mismatches: 96  
Query Match: 10.57% Indels: 45  
DB: 4 Gaps: 7

US-10-631-958-2 (1-326) x US-09-614-221A-399 (1-1875)

Qy 2 LysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyr 21  
Db 679 AGATCGATATTAGTCATTATTAAATCCCAACGGTGTAAAGTACTGCTAAAAATTTATTC 738  
Qy 22 GluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGlyAsn 41  
Db 739 CTGACAAAGCAAGCCCAATACTAGTGGAAGT-----GGCTGC 777  
Qy 42 LysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGluThr 61  
Db 778 AAAATAGAAATTCATAC-----ACAAAATATGCCGTCTACGCCCATCGATATT 825  
Qy 62 LeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyMet 81  
Db 826 GCCAAAGATTAGATATACGAAATACGATACCATTCGATGTGCTCGGTGATGTTATT 885  
Qy 82 PheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGln 101  
Db 886 CCATACGAAGTAATTAATGGCTTTATAGAAGACCCGACAGA-----GTGGATGCG 936  
Qy 102 AsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGly 121  
Db 937 TTCAATAAACATAGCCGTAACTCAG-----CTACCTTGGCGT 972  
Qy 122 SerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeu 141

Db 973 TCAGGAATGCTATGAGCATTTTCATGTTCATTGGACAAATAACCCATCGTACGCGCTCTG 1032  
Qy 142 HisIleValValGlyAspSerLeuAlaMetAspVal-----SerSerValHisHis 158  
Db 1033 TGCGTTGTCAAATCCATTGAAACAAGAATAGACTAATGTGTGTTCCCAACCTTCTTAT 1092  
Qy 159 AsnSerThrLeuLeuArgTyrSerValSerLeuGlyTyrGlyPheTyrGlyAspIle 178  
Db 1093 ATGAACGAATGGCCAAGATTATCTTTTTCAGTCAGACGTCACGGCGTTATTGCAGAATCT 1152  
Qy 179 IleLysAspSerGluLysLysArgTTPLeuGlyLeuAlaArgTyrAspPheSerGlyLeu 198  
Db 1153 GATATTAAACATCGAATTCATCAGATGGTCCCGTCCCGTTAGTTTAATTTGGGT----- 1206  
Qy 199 LysThrPheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis 218  
Db 1207 -----GTAGCATTCACATATTCCAAGT----- 1230  
Qy 219 ThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGln 238  
Db 1231 -----AGAAATATCCCTGCGAAGTTTTCGTCAAATATGCTGCCAAA 1272  
Qy 239 SerLysGlnGlnLeuGlu 244  
Db 1273 TCTAAAAAGGAATTAATA 1290

## RESULT 14

US-09-620-312D-796  
; Sequence 796, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yungqing  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghast  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt\_FL\_genes Version 1.0  
; SEQ ID NO 796  
; LENGTH: 2462  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (196)..(1464)  
US-09-620-312D-796

Alignment Scores:  
Pred. No.: 4.85e-12 Length: 2462  
Score: 173.50 Matches: 62  
Percent Similarity: 45.38% Conservative: 46

```
Best Local Similarity: 26.05% Mismatches: 83
Query Match: 10.10% Indels: 47
DB: 4 Gaps: 11

US-10-631-958-2 (1-326) x US-09-620-312D-796 (1-2462)

QY 2 LysHisLeuLeuValPheIleAsnProPheGlyGlyLeuGlyGlnGlyLysArgIleTyr 21
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
376 AAGAGGCCACTGTTTCTCATCTCGAGCTTGCNAGGAAAGCCAGGACTCTATTT 435

QY 22 GluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleleGlyAsn 41
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
436 GAAAAAATGCTGCCCGATTTTACATTTATCTGCGCATGGATGTCATTT 486

QY 42 LysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGluThr 61
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
487 -----GTTAAGACAGATTATGAGGGCAAGCCAAAGAAA--- 519

QY 62 LeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyValAspGlyMet 81
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
520 CTCCTGGAACG---ATGGAAACACGGATGTGATCATTTGTCAGGAGGATGGACA 576

QY 82 PheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGln 101
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
577 CTGACGAGGTGTTACTGTCTCTCGACGACAGATGAGGCTACCTTC----- 627

QY 102 AsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleleProAlaGly 121
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
628 -----AGTAAGATTCCCATTTGGATTTATCCCACTGGGA 660

QY 122 SerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThr----- 138
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
661 GAGACAGTAGTTGAGTCATACCTCTTTGCGAAGTGGAAACAAAGTCCAAATATT 720

QY 139 -----SerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerVal 156
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
721 ACTGATGCCACACTTGCCATTCTGAAAGGAGAGACAGTTCACCTTGATGTCTTGCAATC 780

QY 157 HisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGly 176
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
781 AAGGGTGAAGAGAACAGACCTGTATTGCAATGACCGGCTTCGATGGGATCTTTTCA 840

QY 177 Asplle---IleLysAspSerGluLysLysArgTyrP-----LeuGlyLeu--- 190
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
841 GATGCTGGCGTCAAGATTAGC-----AAGTACTGGTATCTTGAGCCCTCTAAATAACAA 894

QY 191 AlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThr 210
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
895 GCAGCCCACTTTTTCAGCACTCTTAAGGAGTGGCTCAGACTCAT-----CAAGCCTCT 948

QY 211 ValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysPro 228
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
949 ATCTCATACG-----GGACCTTACAGAGAGACCTCCC 981

RESULT 15
US-09-270-767-30448
; Sequence 30448, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30448
; LENGTH: 901
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-30448

Alignment Scores:

Pred. No.: 1.28e-12 Length: 901
Score: 172.00 Matches: 50
Percent Similarity: 38.59% Conservative: 21
Best Local Similarity: 27.17% Mismatches: 45
Query Match: 10.02% Indels: 68
DB: 4 Gaps: 5

US-10-631-958-2 (1-326) x US-09-270-767-30448 (1-901)

QY 199 LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeu----- 214
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2 AAGGCTTCTCTGAATAATCGCGCTATGACGCCGAATCTGAGAAATCTTAGAAGAGCCGAT 61

QY 215 -----ProAlaGlnHisThrValGlySerProArgAsp----- 225
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
62 CTTTACTGACACCGCGCTGAGGACATTCGCGAGATCCGGATAGTGTGTGCTCGCTG 121

QY 226 -----ArgLysProCysArgAlaGlyCysPheValCys----- 236
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
122 GGAGAGTCAGTGCCATCCGTCTGTATGCCAATTCGCCAGCGCTGCGAGCTTCGCCAGCAGC 181

QY 237 -----ArgGlnSerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGly 253
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
182 ATACAGGAACAGCGATCCTCATTTGTTTCATCCAAGAGGAATCTAAAGAGGCA----- 232

QY 254 LeuGluAlaAlaGluAspValGlu----- 261
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
233 ---GAGCCCAATCAGCAGGTAGAACAGAGGACTCTCATCTAGCCGCCAGTGAAGACGA 289

QY 261 ----- 261
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
290 CTCCTGAGGCCTCGTCCGCTCCAGGCAATCTTCGATTGCCCACTGGCTCCATTTTCATCA 349

QY 262 -----GluTrpGlnValValCysGlyLysPheLeuAlaIleAsn 274
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
350 ATGAGGAACCTCGGCAACGATCAGTGGAAAGTTGTGCGGGCAATTTCTTTATGATCTGC 409

QY 275 AlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaIleHis 294
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
410 GCGCGAACATAACCTGCGCTGCGCCAGGAGTCCCAATGGCATCTCCCGTTACAGTCAT 469

QY 295 LeuGlyAspGlySerSerAspLeuIleLeuLysCysSerArgPheAsnPheLeu 314
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
470 CTGGTGATGGTTGCTGGACCTGATTCTCGTGAAGAAAGACCTCACTGCTCAACACGCTG 529

QY 315 ArgPheLeuIle 318
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
530 CGTTTCTGCTC 541

Search completed: September 6, 2005, 02:33:28
Job time : 208.39 secs
```

**This Page Blank (uspto)**



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: September 5, 2005, 23:41:54 ; Search time 918.291 Seconds  
(without alignments)  
2324.913 Million cell updates/sec

Title: US-10-631-958-2  
Perfect score: 1717  
Sequence: 1 PKHLLVFNPFGKQGGKRI.....KCSRPNFLRLRHTNQDQ 326

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 7338684 seqs, 3274456166 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-DB=US10-631-958-2 -O=FASTA -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPEXT=0 -UNIT8=bits -START=1 -END=1 -MATRIX=blomsum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10631958 @CGN 1 1588 @runat 02092005 165813 4470  
-NCPU=6 -ICPU=3 -NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq.\*  
22: /cgn2\_6/ptodata/1/pubpna/US10J\_NEW\_PUB.seq.\*  
23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq.\*  
24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1717	100.0	979	10	US-09-969-896-1
2	1717	100.0	979	19	US-10-631-958-1
3	1717	100.0	1840	9	US-09-784-810A-5
4	1717	100.0	1840	22	US-10-876-281-5
5	1640.5	95.5	1614	10	US-09-969-896-9
6	1640.5	95.5	1614	19	US-10-631-958-9
7	1640.5	95.5	1740	18	US-10-262-511-39
8	1640.5	95.5	4413	10	US-09-969-896-16
9	1640.5	95.5	4413	19	US-10-631-958-16
10	1640.5	95.5	4429	19	US-10-618-941-55
11	1633.5	95.1	4432	17	US-10-120-988-148
12	1632.5	95.1	4463	16	US-10-315-597A-1
13	811	47.2	474	10	US-09-969-896-4
14	811	47.2	474	19	US-10-631-958-4
15	529	30.8	2657	20	US-10-425-115-176413
16	508.5	29.6	329	10	US-09-969-896-5
17	508.5	29.6	329	19	US-10-631-958-5
18	432.5	25.2	2079	18	US-10-425-114-34530
19	432.5	25.2	2079	20	US-10-425-115-86211
20	414.5	24.1	564	22	US-10-477-445-46
21	335	19.5	1833	19	US-10-437-963-30650
22	334.5	19.5	2166	20	US-10-425-115-64890
23	333	19.4	2084	18	US-10-424-599-47396
24	333	19.4	2174	18	US-10-425-114-34081
25	330	19.2	1869	18	US-10-425-114-34806
26	321.5	18.7	2189	19	US-10-437-963-92801
27	318	18.5	3362	20	US-10-425-115-29002
28	311.5	18.1	1512	20	US-10-425-115-64892
29	308.5	18.0	2014	18	US-10-424-599-59419
30	305.5	17.8	1857	18	US-10-425-114-32139
31	305.5	17.8	2247	20	US-10-425-115-97849
32	302	17.6	167	10	US-09-969-896-6
33	302	17.6	167	19	US-10-631-958-6
34	300.5	17.5	1817	19	US-10-115-635-247
35	300	17.5	1386	19	US-10-437-963-95148
36	291	16.9	2629	19	US-10-437-963-32906
37	289.5	16.9	1468	19	US-10-767-701-13363
38	288.5	16.8	2284	20	US-10-425-115-79424
39	287.5	16.7	3008	19	US-10-437-963-77896
40	282	16.4	2629	17	US-10-348-052-24
41	282	16.4	2629	19	US-10-622-011-24
42	276	16.1	861	20	US-10-425-115-86219
43	273	15.9	522	9	US-09-784-810A-7
44	273	15.9	522	22	US-10-876-281-7
45	270.5	15.8	2609	17	US-10-348-052-25

ALIGNMENTS

RESULT 1  
US-09-969-896-1  
; Sequence 1, Application US/09969896  
; Publication No. US20030125533A1  
; GENERAL INFORMATION:  
; APPLICANT: Kossida, Sophia  
; TITLE OF INVENTION: Regulation of human Sphingosine  
; FILE REFERENCE: 004974.00594  
; CURRENT APPLICATION NUMBER: US/09/969,896  
; CURRENT FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: US 60/238,005  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: US 60/314,113  
; PRIOR FILING DATE: 2001-08-23  
; NUMBER OF SEQ ID NOS: 16

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 979
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-1

Alignment Scores:
Pred. No.: 1.24e-211 Length: 979
Score: 1717.00 Matches: 326
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-10-631-958-2 (1-326) x US-09-969-896-1 (1-979)

Qy 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
Db 2 CCAAGCATTTACTGGTATTTATCAACCCGTTTGGAGGAAAGGACCAAGGCAAGCGGATA 61
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40
Db 62 TATGAAAGAAAGTGGCCACCACTGTTACCTTAGCCTCCATCACCCTGACATCATCGGT 121
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 122 AACAAATCTATGTTAACTATGATAGAGTAATTAATCACTACCTCCATCAGCCCAAGGAG 181
Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
Db 182 ACTCTGATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTCTGGCGGAGATGGT 241
Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 242 ATGTTTCAGCAGGTGCTGCAACCCGTTTGGAGGAGCGCAGAGGAGCGCGGGTGCAC 301
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db 722 CAGCAGCTGGAGGAGGACAGAGAAACACTGTTATGGTTTGGAAAGCTGCGGAGGACGTG 781
Qy 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 782 GAGGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACCAACATGTCCTGT 841

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 979
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-958-1

Alignment Scores:
Pred. No.: 1.24e-211 Length: 979
Score: 1717.00 Matches: 326
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 19 Gaps: 0

US-10-631-958-2 (1-326) x US-10-631-958-1 (1-979)

Qy 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
Db 2 CCAAGCATTTACTGGTATTTATCAACCCGTTTGGAGGAAAGGACCAAGGCAAGCGGATA 61
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40
Db 62 TATGAAAGAAAGTGGCCACCACTGTTACCTTAGCCTCCATCACCCTGACATCATCGGT 121
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 122 AACAAATCTATGTTAACTATGATAGAGTAATTAATCACTACCTCCATCAGCCCAAGGAG 181
Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
Db 182 ACTCTGATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTCTGGCGGAGATGGT 241
Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 242 ATGTTTCAGCAGGTGCTGCAACCCGTTTGGAGGAGCGCAGAGGAGCGCGGGTGCAC 301
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db 302 CAGAACCAACCCCGGGCTGCTGCTGCCAGTAGCCTCCGATTTGAATCATTTCCGCA 361
Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 362 GGGTCAACGGACTGCGTGTGTTACTCCACCGTGGCCACGACGACGAGAAACCTCGCG 421
Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisAsnSer 160
Db 422 CTGCATATCGTGTGTTGGGACTCGCTGGCCATGGATGTCTCCTCAGTCCACCAACAGC 481
Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db 482 ACATCTCTCTACTCGCTGCTCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAG 541
Qy 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 542 GACAGTGAGAGAAACCGGTGCTGGCTCTGCCAGATACGACTTTTCAGGTTTAAAGACC 601
Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 602 TTCCTCTCCCAACCATGCTATGAAAGGACAGTGTCTTCCTCCCTGCAACACACCGGTG 661
Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db 662 GGATCTCAAGGATAGAAAGCCCTGCGCGCAGGATGCTTTGTTTTCAGGCAAGCAAG 721
Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 722 CAGCAGCTGGAGGAGGACAGAGAAACACTGTTATGGTTTGGAAAGCTGCGGAGGACGTG 781
Qy 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 782 GAGGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACCAACATGTCCTGT 841
```

Db 362 GGGTCAACGACTGCGTGTGTTACTCCACCGTGGGACCAGCGACGAGAACTTCGGCG 421  
Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160  
Db 422 CTGCATATCGTGTGTTGGGACTCGCTGGCGATGATGTCTCAGTCCACCACACAGC 481  
Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180  
Db 482 ACACATCCCTTCGCTACTCCGCTGTCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAG 541  
Qy 181 AspSerGluLysLeuArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200  
Db 542 GACAGTGAGAGAAACCGTGTGTTGGTCTTCGCAGATACGACTTTTCAGTTTAAAGACC 601  
Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220  
Db 602 TTCCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCCTCCCTGCACAAACACACGCTG 661  
Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240  
Db 662 GGATCTCCAAAGGATAGGAAGCCCTGCGCGGAGGATGCTTTGTTTGGAGGCAAGCAAG 721  
Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260  
Db 722 CAGCAGCTGGAGGAGGAGCAGAGAAAGCACTGTATGGTTTGGAGCTGGGAGACGCTG 781  
Qy 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280  
Db 782 GAGGAGTGGCAAGTCTGTCTGGGAAGTTTCTGGCCATCAATGTCACAAACATGTCTGT 841  
Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300  
Db 842 GCTTGTGCGCGAGAGCCCAAGGGCCCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCT 901  
Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320  
Db 902 GACCTCATCTCATCCGGAATGCTCCAGGTTCATATTTCTGAGATTTCTCATCAGGCAC 961

## RESULT 3

US-09-784-810A-5  
; Sequence 5, Application US/09784810A  
; Patent No. US20020082203A1

## GENERAL INFORMATION:

; APPLICANT: RASTELLI, LUCA  
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING  
; TITLE OF INVENTION: SAME  
; FILE REFERENCE: 10716-08  
; CURRENT APPLICATION NUMBER: US/09/784,810A  
; CURRENT FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: 60/182,360  
; PRIOR FILING DATE: 2000-02-14  
; PRIOR APPLICATION NUMBER: 60/191,261  
; PRIOR FILING DATE: 2000-03-22  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 5  
; LENGTH: 1840

## ; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-784-810A-5

## Alignment Scores:

Pred. No.: 3,31e-211 Length: 1840  
Score: 1717.00 Matches: 326  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

## RESULT 4

US-10-876-281-5

; Sequence 5, Application US/10876281

US-10-631-958-2 (1-326) x US-09-784-810A-5 (1-1840)

Qy 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20  
Db 265 CCAAAGCATTTACTGGTATTTATCAACCCGTTTGGAGGAAAGGACAAGGCAAGCGGATA 324  
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40  
Db 325 TATGAAAGAAAGTGGCACCACTGTTACCTTAGCCTCCATCACCACATGACATCATCGGT 384  
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60  
Db 385 AACAAATTTCTATGTTAACTATGTAGAGTATTACTGAACATGCTTAATCAGCCAGGAG 444  
Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80  
Db 445 ACTCTGTATGAGATTAAACATAGACAAATACACGCGCATCGTCTGTGTCGGGGAGATCGT 504  
Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100  
Db 505 ATGTTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGCGCGGGGTTCGAC 564  
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120  
Db 565 CAGAACCAACCCCGGGCTGTCTGGTCCCAAGTAGCCTCCGGATTGGAAATCAATCCCGCA 624  
Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140  
Db 625 GGGTCAACGGACTGCGTGTGTTACTTCCACCGTGGSCACACGACGACGAGAACTTCGGCG 684  
Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160  
Db 685 CTGATATTCGTTGTTGGGGACTCGCTGGCCATGGAATGTCTCTCAGTCCACACACAGC 744  
Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180  
Db 745 ACATCTCTTCCTACTCGTGTCTCCTGCTGGGCTACGGCTTCTACGGGGAGCATCATCAAG 804  
Qy 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200  
Db 805 GACAGTGAGAGAAACCGGTGCTTGGCTCTTCCAGATACGACTTTTCAAGTTTAAAGACC 864  
Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220  
Db 865 TTCCTCTCCCACTGCTATGAAGGGACAGTGTCTTCTCCTCCCTGCACACACACCGGTG 924  
Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240  
Db 925 GGATCTCCAAAGGATAGGAAGCCCTGCGCGGACGAGATGCTTTGTTTGGAGGACGAG 984  
Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260  
Db 985 CAGCAGCTGGAGGAGGAGCAGAGAAAGCACTGTATGTTTGGAAAGCTGCGGAGACGCTG 1044  
Qy 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280  
Db 1045 GAGGAGTGGCAAGTCTGTGTTGGGAAGTTTCTGCCCATCAATGCCACAAACATGTCTGT 1104  
Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300  
Db 1105 GCTTGTGCGCGAGGCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCT 1164  
Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320  
Db 1165 GACCTCATCTCATCCGGAATGCTCCAGGTTCATATTTTCTGAGATTTCTCATCAGGCAC 1224

Qy 321 ThrAsnGlnGlnAspGln 326

Db 1225 ACCAACCAAGCAGGACCAG 1242

```
; Publication No. US20050123942A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: SAME
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/10/876,281
; CURRENT FILING DATE: 2004-06-24
; PRIOR APPLICATION NUMBER: US/09/784,810
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1840
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-876-281-5

Alignment Scores:
Pred. No.: 3,31e-211 Length: 1840
Score: 1717.00 Matches: 326
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-10-631-958-2 (1-326) x US-10-876-281-5 (1-1840)

Qy 1 ProlystHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgile 20
Db 265 CCAAAGCATTTACTGGTATTTATCAACCGTTTGGAGAAAGCAAGCAAGCGGATA 324

Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleGly 40
Db 325 TATGAAGAAAGTGGCACCCTGTTACCTTAGCTTACCTCCATCACCCTGATCATCGGT 384

Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 385 AACAAATTTCTATGTTAATATGTAGAAAGTAAATTAATGTAACATGCTAATCAAGCGCAAG 444

Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyLysValGlyGlyAspGly 80
Db 445 ACTGTATGATGATTAACATAGACAAATACACCGCATCTGCTGTGCGCGGAGATGGT 504

Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 505 ATGTTTCAGCGAGGTGCTGCACGGTCTGATTGGAGGACGACAGAGCGCGCGGTCGAC 564

Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db 565 CAGAACCAACCCCGGGCTGTGCTGTCGCCAGTAGCTCCGGATTGGAATCATTTCCGCA 624

Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 625 GGGTCAACGGACTGGGTGTGTACTTCCACCGTGGCGACCGACGACGAGAAACCTCGCG 684

Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
Db 685 CTGCATATCGTTGTTGGGACTCGCTGGCCATGATGTGTCCTCAGTCCACCAACACAGC 744

Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db 745 ACATCTCTTCGCTACTCCGCTGCTCCCTGCTGGCTACGGCTTCTACGGGGAATCATCAAG 804

Qy 181 AspSerGluLysValArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 805 GACAGTGAGAGAAACGGTGGTGTGCTTCCAGATACGACTTTTCAGGTTTAAAGACC 864

Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220

; Publication No. US20050123942A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: SAME
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/10/876,281
; CURRENT FILING DATE: 2004-06-24
; PRIOR APPLICATION NUMBER: US/09/784,810
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1840
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-876-281-5

Alignment Scores:
Pred. No.: 3,31e-211 Length: 1840
Score: 1717.00 Matches: 326
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-10-631-958-2 (1-326) x US-09-969-896-9 (1-1614)

Qy 1 ProlystHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgile 20
Db 388 CCAAAGCATTTACTGGTATTTATCAACCGTTTGGAGAAAGCAAGCAAGCGGATA 447

Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40
Db 448 TATGAAGAAAGTGGCACCCTGTTACCTTACCTTCCATCACCCTGATCATCATC 504

Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 505 -----GTTACTGAAACATGCTATCATCGGCCAAGGAG 534

Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
```

```

Db      535  ACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGGT 594
Qy      81   MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db      595  ATGTTTCAGGAGGTGTCGACGGTCTGTGATTGGGAGGACGACAGAGGAGCGCGGGTTCGAC 654
Qy      101  GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db      655  CAGAACACCCCGGGCTGTGTCGCCAGTAGCTCCGGATTGGAAATCATTCCTCCGCA 714
Qy      121  GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db      715  GGGTCAACGGACTGCGTGTGTTACTCCACCGTGGGACCCAGCGACGACGAGAACTTCGGG 774
Qy      141  LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
Db      775  CTGCATATCTGTTGGGACATCGCTGGCCATGGATGTGCTCCTAGTCCACCACACAGC 834
Qy      161  ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db      835  ACACCTCTTCGCTACTCCCGTGTCTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATCA 894
Qy      181  AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db      895  GACAGTGAGAGAAACGGTGTGGTCTTGGCTTGTGCATACGACTTTTCAGGTTTAAAGACC 954
Qy      201  PheLeuSerHisGlyCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db      955  TTCTCTCTCCACCATCTGCTATGAGGGACAGTGTCTTCTCTCCCTGCAACACACACGGTG 1014
Qy      221  GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db      1015  GGATCTCAAGGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTGAGGCAAAAGCAAG 1074
Qy      241  GlnGlnLeuGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db      1075  CAGCAGCTGGAGAGAGAGAGAAAGCAGTGTATGGTTTGGAAAGCTGCGGAGGACGTG 1134
Qy      261  GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db      1135  GAGGAGTGGCAAGTCTGCTGCGGAAGTTCAGAGTTCAATTTCTGAGATTTCATCAGGCAC 1314
Qy      301  AspLeuIleLeuLeuArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
Db      1255  GACCTCATCTCTCATCCGAAATGCTCCAGGTTCAATTTCTGAGATTTCATCAGGCAC 1314
Qy      321  ThrAsnGlnGlnAspGln 326
Db      1315  ACCAACACGAGGACACAG 1332

```

## RESULT 6

```

US-10-631-958-9
; Sequence 9, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16

```

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 9

; LENGTH: 1614

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-631-958-9

## Alignment Scores:

```

Pred. No.:      2,22e-201      Length:      1614
Score:          1640.50      Matches:      314
Percent Similarity: 96.63%      Conservative: 1
Best Local Similarity: 96.32%      Mismatches: 0
Query Match:     95.54%      Indels:      11
Db:              19          Gaps:       1

```

US-10-631-958-2 (1-326) x US-10-631-958-9 (1-1614)

```

Qy      1   ProlHisHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
Db      388  CCAAGCAATTTACTGTTATTTATCACCCTTTGGAGGAAAGGACAGGACGAGGATA 447
Qy      21   TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
Db      448  TATGAAAGAAAGTGGACCACTGTTCACTTAGCTTCATCACCACCTGACATCATC--- 504
Qy      41   AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db      505  -----GTTACTGAACATGCTTAATCAGGCCAAGGAG 534
Qy      61   ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
Db      535  ACTCTGTATGAGATTAACTAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGGT 594
Qy      81   MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db      595  ATGTTTCAGGAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGGAGCGCGGGTTCGAC 654
Qy      101  GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db      655  CAGAACACCCCGGGCTGTGTCGCCAGTAGCTCCGGATTGGAAATCATTCCTCCGCA 714
Qy      121  GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db      715  GGGTCAACGGACTGCGTGTGTTACTCCACCGTGGGACACGACGACGAGAACTTCGGG 774
Qy      141  LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
Db      775  CTGCATATCTGTTGGGACTTCGCTGGCCATGGATGTGCTCCTAGTCCACCACACAGC 834
Qy      161  ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db      835  ACACCTCTTCGCTACTCCCGTGTCTCCAGGTCAGGCTTCTACGGGGACATCATCAAG 894
Qy      181  AspSerGluLysLysArgTyrPheLeuGluAlaArgTyrAspPheSerGlyLeuLysThr 200
Db      895  GACAGTGAGAGAAACGGTGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAGACC 954
Qy      201  PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db      955  TTCTCTCTCCACCATCTGCTATGAGGGACAGTGTCTTCTCTCCCTGCAACACACAGG 1014
Qy      221  GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db      1015  GGATCTCAAGGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTGAGGCAAAAGCAAG 1074
Qy      241  GlnGlnLeuGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db      1075  CAGCAGCTGGAGGAGGACGACAGAAAGCAGTGTATGGTTTGGAAAGCTGCGGAGGACGTG 1134
Qy      261  GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db      1135  GAGGAGTGGCAAGTCTGCTGCGGAAGTTCAGGTTCAATTTCTGAGATTTCATCAGGCAT 1194

```

Qy	281	AlaCysArgArgSerProArgClyLeuSerProIalaHisLeuGlyAspClySerSer	300
Db	1195	GCITGTGCGCGAGCCCGAGGGGCTCTCCCGCGCTGCGCCACTTGGGAGACGGGTCTTCT	1254
Qy	301	AspLeuLeuLeuLeuArgLysCysSerArgPheAsnPheLeuArgPheLeuLeuArgHis	320
Db	1255	GACCTCATCTCTCATCCGGAAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCAC	1314
Qy	321	ThrAsnGlnGlnAspGln	326
Db	1315	ACCAACCCAGCAGGACCCAG	1332

## RESULT 7

```

US-10-262-511-39
; Sequence 39, Application US/10262511
; Publication No. US20040038223A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glennnda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zehusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraSeqList version 0.1

```



Db 1210 GAGGAGTGGCAAGTCTGCTGTGGGAAGTTCTGGCCATCAATGCCCAACATGTCCTGT 1269  
Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300  
Db 1270 GCTTGTGCGCGGAGGCCAGGCCCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTCT 1329  
Qy 301 AspLeuLeuLeuLeuArgLysCysSerArgPheAsnPheLeuArgPheLeuLeuLeuHis 320  
Db 1330 GACCTCATCTCTATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTCTCATCAGGCAC 1389  
Qy 321 ThrAsnGlnGlnAspGln 326  
Db 1390 ACCAACCCAGCAGGACCAG 1407

## RESULT 8

US-09-969-896-16  
; Sequence 16, Application US/09969896  
; Publication No. US20030125533A1  
; GENERAL INFORMATION:  
; APPLICANT: Kossida, Sophia  
; TITLE OF INVENTION: Regulation of human Sphingosine  
; FILE REFERENCE: Kinase-Like Protein  
; CURRENT APPLICATION NUMBER: US/09/969,896  
; CURRENT FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: US 60/238,005  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: US 60/314,113  
; PRIOR FILING DATE: 2001-08-23  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-969-896-16

Alignment Scores:  
Pred. No.: 1.07e-200 Length: 4413  
Score: 1640.50 Matches: 314  
Percent Similarity: 96.63% Conservative: 1  
Best Local Similarity: 96.32% Mismatches: 0  
Query Match: 95.54% Indels: 11  
Db: 10 Gaps: 1

US-10-631-958-2 (1-326) x US-09-969-896-16 (1-4413)

Qy 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLeuGlyGlnGlyLeuArgIle 20  
Db 463 CCNAAAGCATTTACTGGTATTTATCAACCGTTTGGAGGAAAGGCAAGGCGGATA 522  
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40  
Db 523 TATGAAAGAAAGTGGCACCAGTTCACCTTAGCTTACCTCACCACCTGACATCATC--- 579  
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60  
Db 580 -----GTTACTGAACATGCTAAATCAGGCCCAAGGAG 609  
Qy 61 ThrLeuTyrGluLeuAsnIleAspLysTyrAspGlyIleValCysValGlyAspGly 80  
Db 610 ACTGTGTAGATTAAACATACAAAATACGACGCGCATCGTGTGTGGCGGAGATGGT 669  
Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100  
Db 670 ATGTTCCAGGAGGTCTGCAGGCTGTGATTGGAGGAGCGCAGGAGCGCGGGTCCAC 729  
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120  
Db 730 CAGAACCACCCCGGGCTGTGCTGGTCCCGAGTAGCTCCGATTGGAATCATTCGCCGA 789  
Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140

Db 790 GGGTCAACGGACTGGTGTGTATTCTCCACCGTGGGCACACGAGCGCAGAAACCTCGGG 849  
Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160  
Db 850 CTGCATATCGTTGTTGGGCACTCGCTGGCCATGGATGTGTCTCAGTCCACCAACAGC 909  
Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180  
Db 910 ACATCTCTTCTACTCTCGTGTCTGCTGGGCTACGGCTTCTACGGGGGACATCATCAAG 969  
Qy 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200  
Db 970 GACAGTGAGAGAAACGGTGGTGGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGACC 1029  
Qy 201 PheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220  
Db 1030 TTCCTCTCCCACTGCTATGAAGGACAGTGTCTTCTCCCTGCGACAAACACACAGGTG 1089  
Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240  
Db 1090 GGATCTCCAAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTGCAGGCAAGCAAG 1149  
Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260  
Db 1150 CAGCAGCTGGAGGAGGACAGCAAGAACACTGTATGGTTTGGAAAGCTGCGGAGACGTG 1209  
Qy 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280  
Db 1210 GAGGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCTGT 1269  
Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300  
Db 1270 GCTTGTGCGCGGAGCCCGCAGGGGCTCTCTCCCGGCTGCCACCTTGGGAGACGGGTCTTCT 1329  
Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320  
Db 1330 GACCTCATCTCTATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTCTCATCAGGCAC 1389

## RESULT 9

US-10-631-958-16  
; Sequence 16, Application US/10631958  
; Publication No. US20040192580A1  
; GENERAL INFORMATION:  
; APPLICANT: Kossida, Sophia  
; TITLE OF INVENTION: Regulation of human Sphingosine  
; FILE REFERENCE: Kinase-Like Protein  
; CURRENT APPLICATION NUMBER: US/10/631,958  
; CURRENT FILING DATE: 2003-08-01  
; PRIOR APPLICATION NUMBER: US/09/969,896  
; PRIOR FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: US 60/238,005  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: US 60/314,113  
; PRIOR FILING DATE: 2001-08-23  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 4413  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-631-958-16

Alignment Scores:  
Pred. No.: 1.07e-200 Length: 4413  
Score: 1640.50 Matches: 314  
Percent Similarity: 96.63% Conservative: 1  
Best Local Similarity: 96.32% Mismatches: 0  
Query Match: 95.54% Indels: 11



```
DB: 19 Gaps: 1
US-10-631-958-2 (1-326) x US-10-631-958-16 (1-4413)
QY 1 ProlysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
DB 463 CCAAAGCATTTACTGGTATTTATCAACCGTTTGGAGGAAAAGGACAAGGCAAGCGGATA 522
QY 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
DB 523 TATGAAAGAAAGTGGCAACTGTTTCACCTTAGCTTCATCAGCTCATCATCATC--- 579
QY 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
DB 580 -----GTTACTGAACATGCTAATCAGGCCCAAGGAG 609
QY 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
DB 610 ACTCTGTATGAGATTAAACATAGACAAATACAGCGCATCGTCTGTGTGGCGGAGATGGT 669
QY 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
DB 670 ATGTTACAGAGGTGCTCAGCGTCTGATTGGAGAGCGCAGAGAGCGCGGGTGCAC 729
QY 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
DB 730 CAGAACCAACCCCGGGCTGTGCTGTCGCCAGTAGTACGCTTCGGATTGGGAATCATTCGCCA 789
QY 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
DB 790 GGGTCAACGGGACTGCGTGTGTACTCCACCGTGGGCACAGCGAGAGAACTTCGGCG 849
QY 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
DB 850 CTGCATATCGTTTGGGACTCGCTGGCCATGATGTCTCAGTCCACCAACACACG 909
QY 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
DB 910 ACACCTCTTCGCTACTCCGCTGTCCTGCTGGCTACGGCTTCTACGGGGACATCATCAAG 969
QY 181 AspSerGluLysLysArgTyrLeuGlyThrLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
DB 970 GACAGTGAGAAAGAACCGTGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAGACC 1029
QY 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
DB 1030 TTCTCTCCCAACCACTGCTATGAAGGACAGTGTCTTCTCCCTCCCTGCACACACACGGTG 1089
QY 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
DB 1090 GGATCTCAAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTGTCAGGCCAAGCAAG 1149
QY 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
DB 1150 CAGCAGCTGGAGGAGGAGCAGAAAGCACTGTATGGTTTGGAGCTGCGGAGACGCTG 1209
QY 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
DB 1210 GAGGAGTGCGAAGTCTGTCTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCCTGT 1269
QY 281 AlaCysArgArgSerProArgGlyLeuSerProAlaHisLeuGlyAspGlySerSer 300
DB 1270 GCTTGTCCCGAGAGCCCGAGGGGCTCTCCCGGGCTGCCACTTGGGAGACGGGTCTTCT 1329
QY 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
DB 1330 GACCTCATCTCATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAC 1389
QY 321 ThrAsnGlnGlnAspGln 326
DB 1390 ACCAACCAAGCAGGACCAG 1407
```

RESULT 10

```
US-10-618-941-55
; Sequence 55, Application US/10618941
; Publication No. US20040197792A1
; GENERAL INFORMATION:
; APPLICANT: WHITE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: CAENEPEEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618,941
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 55
; LENGTH: 4429
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-618-941-55
Alignment Scores:
Pred. No.: 1.07e-200 Length: 4429
Score: 1640.50 Matches: 314
Percent Similarity: 96.63% Conservative: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 95.54% Indels: 11
DB: 19 Gaps: 1
US-10-631-958-2 (1-326) x US-10-618-941-55 (1-4429)
QY 1 ProlysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
DB 479 CCAAAGCATTTACTGGTATTTATCAACCGTTTGGAGGAAAAGGACAAGGCAAGCGGATA 538
QY 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
DB 539 TATGAAAGAAAGTGGCACCACCTGTTTCCCTTACCTTCCATCCACTGCATCATCATC--- 595
QY 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
DB 596 -----GTTACTGAACATGCTAATCAGGCCCAAGGAG 625
QY 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
DB 626 ACTCTGTATGAGATTAAACATAGACAAATACAGCGCATCGTCTGTGTGGCGGAGATGGT 685
QY 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
DB 686 ATGTTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGGAGCGCGGGGTGCAC 745
QY 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
DB 746 CAGAACCAACCCCGGGCTGTGTCCTCCAGTAGCTCCCGGATTTGGAATCATTTCCCGCA 805
QY 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
DB 806 GGGTCAACGGACTCGGTGTGTACTCCACCGTGGGACCGACGACGAGAAACCTCGGCG 865
QY 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
DB 866 CTGCATATCGTGTGTGGGACTCGCTGGCCATGATGTGTCTCAGTCCACACACACAGC 925
QY 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
DB 926 ACATCTCTTCCCTACTCCGTTGTCCTGTGGGCTACGGCTTCTACGGGGACATCATCAAG 985
QY 181 AspSerGluLysLysArgTyrLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
DB 986 GACAGTGAGAAAGCGTGGTGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGACC 1045
QY 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
```

Db 1046 TTCTCTCCACCACCTGCTATGAGGGACAGTGTCTTCTCTCCCTGACACACACAGCGTG 1105  
QY 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240  
Db 1106 GGATCTCCAAGGATAGGAAAGCCCTCCCGGCGAGGATGCTTTGTTTGCAAGGCAAGCAAG 1165  
QY 241 GlnGlnLeuGluGluGlnLysLysValLeuTyrGlyLeuGluAlaAlaGluAspVal 260  
Db 1166 CAGCAGCTGGAGAGGAGCAGAAAGACACTGTATGTTTGGAAAGCTGCGAGGACGTG 1225  
QY 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280  
Db 1226 GAGGAGTGCGCAAGTGTCTGTGGAGATTCTTGCCCATCAATGCCACCAACATATGCTCTGT 1285  
QY 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300  
Db 1286 GCTTGTGCGCGAGGCCCGGAGGCGCTCTCCCGGCTGCCCACTTGGAGAGCGGCTTTCT 1345  
QY 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320  
Db 1346 GACCTCATCTCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATAGGCAC 1405  
QY 321 ThrAsnGlnGlnAspGln 326  
Db 1406 ACCAACCAGCAGGACCAG 1423

## RESULT 11

US-10-120-988-148  
; Sequence 148, Application US/10120988  
; Publication No. US20030219745A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Goodrich, Kyle  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Dmanac, Radoje T.  
; TITLE OF INVENTION: No. US20030219745A1el Nucleic Acids and  
; FILE REFERENCE: 802CON  
; CURRENT APPLICATION NUMBER: US/10/120,988  
; CURRENT FILING DATE: 2002-04-11  
; PRIOR APPLICATION NUMBER: 09/774,528  
; PRIOR FILING DATE: 2001-01-30  
; NUMBER OF SEQ ID NOS: 441  
; SOFTWARE: pt FL\_genes Version 2.0  
; SEQ ID NO 148  
; LENGTH: 4432  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1129)..(2817)  
US-10-120-988-148

## Alignment Scores:

Pred. No.: 8,686-200 Length: 4432  
Score: 1633.50 Matches: 313  
Percent Similarity: 96.32% Conservative: 1  
Best Local Similarity: 96.01% Mismatches: 1  
Query Match: 95.14% Indels: 11  
DB: 17 Gaps: 1

US-10-631-958-2 (1-326) x US-10-120-988-148 (1-4432)

QY 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20  
Db 1591 CCAAAGCATTTTACTGGTATTTATCAACCCGTTTGAGGAAAAGGCAAGGCAAGCGGATA 1650  
QY 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40  
Db 1651 TATGAAGAAAAGTGGCAACCACTGTTCACTTAGCTTCATCACCACCTGACATCATC--- 1707

QY 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60  
Db 1708 -----GTTTACTGAACATGCTAATCAGGCCAAGGAG 1737  
QY 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80  
Db 1738 ACTCTGTATGAGATTACATAGACAAATACAGCGGCATCGTCTGTGTCTGCGGGAGATGGT 1797  
QY 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100  
Db 1798 ATGTTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGCGCGCGGTGCGAC 1857  
QY 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120  
Db 1858 CAGAACCAACCCCGGGCTGTGCTGTCCTCCCAAGTACCTCCGATTTGGAATCATTTCCCGCA 1917  
QY 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140  
Db 1918 GGGTCAACGGACTCGGTGTGTACTCCACCGTGGGACACGACGACGACGACGACGACGCG 1977  
QY 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160  
Db 1978 CTGCATATCGTTGTGGGACTCGCTGCCCATGATGTCTCCTCAGTCCACCAACACAGC 2037  
QY 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180  
Db 2038 ACCTCTCTCGCTACTCCGTGCTCGGCTACGGCTTCTACGGGGACATCATCAAG 2097  
QY 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200  
Db 2098 GACAGTGAGAAAGAAACCGTGTGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGACC 2157  
QY 201 PheLeuSerHisHisCysTyrTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220  
Db 2158 TTCTCTCTCCCACTGCTATGGAGGGACAGTGTCTTCTCTCTGACACACACACGCGTG 2217  
QY 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240  
Db 2218 GGATCTCCAAGGATAGGAAAGCCCTGCGGGCAGGATGCTTTGTTGCGAGCAAGCAAG 2277  
QY 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260  
Db 2278 CAGCAGCTGGAGGAGGAGCAGAAAGACACTGTATGTTTGGAAAGCTGCGAGGACGTG 2337  
QY 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280  
Db 2338 GAGGAGTGGAAGTCTGCTGTGGGAAGTTTCTGSCCATCAATGCCACCAACATGTCCTGT 2397  
QY 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300  
Db 2398 GCTTGTGCGCGAGCGCCCGGCGCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTCT 2457  
QY 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320  
Db 2458 GACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATAGGCAC 2517  
QY 321 ThrAsnGlnGlnAspGln 326  
Db 2518 ACCAACCAGCAGGACCAG 2535

## RESULT 12

US-10-315-597A-1  
; Sequence 1, Application US/10315597A  
; Publication No. US20030162206A1  
; GENERAL INFORMATION:  
; APPLICANT: Suglura, Masako  
; APPLICANT: Kono, Keita  
; APPLICANT: Kohama, Takafumi  
; TITLE OF INVENTION: Ceramide  
; FILE REFERENCE: 02658CIP/HG  
; CURRENT APPLICATION NUMBER: US/10/315,597A  
; PRIOR FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: JP 2000-178039

```
; PRIOR FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4463
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 3371
; OTHER INFORMATION:
US-10-315-597A-1
Alignment Scores:
Pred. No.: 1.18e-199 Length: 4463
Score: 1632.50 Matches: 312
Percent Similarity: 96.32% Conservative: 2
Best Local Similarity: 95.71% Mismatches: 1
Query Match: 95.08% Indels: 11
DB: 16 Gaps: 1
US-10-631-958-2 (1-326) x US-10-315-597A-1 (1-4463)
Qy 1 ProlystHisLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgile 20
Db 511 CCAAAGCATTTACTGTTATTTATCAACCCGTTTGGAGGAAAGGACAAAGCGGATA 570
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleGly 40
Db 571 TATGAAAGAAAGTGGCCACCACTGTTACCTTTAGCCTCCATCACCTGACATATC--- 627
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 628 -----GTTACTGAACATGCTATCAGGCCAAGAG 657
Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
Db 658 ACTCTGTATGAGATTAACATAGACAAATACGACGGCATCGTCTGTCTGGCGGAGATGTT 717
Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 718 ATGTTTCAGCAGGTGCTGCACGGTCTGATTGGGAGACGACGAGAGCGCGGGTGCAC 777
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAla 120
Db 778 CAGAACCAACCCCGGGCTGCTGTCGCCAGTAGCCTCCGGATTGGAATCATTTCCGCA 837
Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 838 GGTCCAAACGGACTCGTGTGTTACTCCACCCTGGGCACCGACGACGAGAAACCTCGCG 897
Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerValHisHisAsnSer 160
Db 898 CTGCATATCGTGTGTGGGACTCGCTGGCCATGATGTGCTCAGTCCACCAACAAACAGC 957
Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db 958 ACATCTCTTCTACTCGTGTCTCTGCTGGCTACGGCTTCTACGGGGACATCATCAAG 1017
Qy 181 AspSerGluLysLysArgTyrPheLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 1018 GACAGTGAGAGAAACCGTGGTCTTGGCTTCTGGCTTCTACGGGGACATCATCAAG 1077
Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 1078 TTCTCTCTCCCACTGCTATGAGAGGACAGTGTCTTCTCTCCCTGACACCAACACCGGTG 1137
Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db 1138 GGATCTCAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTTGCAGGCAAGCAAG 1197
Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaGluAspVal 260
```

---

```
Db 1198 CACGACGCTGGAGGAGGAGCAGAGAAGCACTGTATGGTTTGAAGCTGCGGAGGACGTG 1257
Qy 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 1258 GAGGAGTGGCAAGTCTGTGTGGAAAGTTTCTGGCCATCATGCCACAAACATGTCTGT 1317
Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
Db 1318 GCTTGTGCGCGGAGCCCGCAGGGGCCCTCTCCCGGCTGCCACCTTGGAGACGGGTCTTCT 1377
Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
Db 1378 GACCTCATCTCTATCCGGAATGCTCCAAGTTCAATTTTCTGAGATTTCTCATCAGGCAC 1437
Qy 321 ThrAsnGlnGlnAspGln 326
Db 1438 ACCAACCCAGCAGGACCAG 1455
RESULT 13
US-09-969-896-4
; Sequence 4, Application US/09969896
; Publication No. US20030125533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-4
Alignment Scores:
Pred. No.: 1.07e-94 Length: 474
Score: 811.00 Matches: 154
Percent Similarity: 99.35% Conservative: 0
Best Local Similarity: 99.35% Mismatches: 1
Query Match: 47.23% Indels: 0
DB: 10 Gaps: 0
US-10-631-958-2 (1-326) x US-09-969-896-4 (1-474)
Qy 80 GlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyVal 99
Db 8 GGTATGTTTACGAGGAGGTGCTGCGGGTCTGATTGGGAGGACGACGAGAGCGCGGGGTC 67
Qy 100 AspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIlePro 119
Db 68 GACCAGAACCAACCCCGGGCTGTGTGTCCTCCAGTAGCTCCCGGATTGGAATCATTTCCC 127
Qy 120 AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer 139
Db 128 GCAGGGGTCAACGGACTCGTGTGTACTCCACCGTGGGCACCGACGACGAGAACCTCG 187
Qy 140 AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerValHisHisAsn 159
Db 188 GCGCTGCATATCGTGTGTGGGACTCGCTGGCCATGGATGTCTCAGTCCACCAACAC 247
Qy 160 SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIle 179
Db 248 AGCACACTCTCTCGTACTCCGTGTCCTGCTGGCTTACGGCTTCTACGGGGACATCATC 307
Qy 180 LysAspSerGlnLysLysArgTyrPheLeuAlaArgTyrAspPheSerGlyLeuLys 199
```

Db 308 AAGGACAGTGAGAAAGACGGTGGTTCCTTGGCCAGATACGACTTTTCAGGTTTAAAG 367  
Qy 200 ThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr 219  
Db 368 ACCTTCCTCCACCACCTGCTATGAAGGACAGTGTCTCTCTCCCTGCACAACACAG 427  
Qy 220 ValGlySerProArgAspArgLysProCysArgAlaGlyCysPhe 234  
Db 428 GTGGGATCTCCAAGGATAGGAAGCCCTGCGGGGCAAGATGCTTT 472  
RESULT 14  
US-10-631-958-4  
; Sequence 4, Application US/10631958  
; Publication No. US20040192580A1  
; GENERAL INFORMATION:  
; APPLICANT: Koestida, Sophia  
; TITLE OF INVENTION: Regulation of human Sphingosine  
; FILE REFERENCE: Kinase-Like Protein  
; FILE REFERENCE: 004974.00594  
; CURRENT APPLICATION NUMBER: US/10/631,958  
; CURRENT FILING DATE: 2003-08-01  
; PRIOR APPLICATION NUMBER: US/09/969,896  
; PRIOR FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: US 60/238,005  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: US 60/314,113  
; PRIOR FILING DATE: 2001-08-23  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-631-958-4

Alignment Scores:  
Pred. No.: 1,07e-94 Length: 474  
Score: 811.00 Matches: 154  
Percent Similarity: 99.35% Conservative: 0  
Best Local Similarity: 99.35% Mismatches: 1  
Query Match: 47.23% Indels: 0  
DB: 19 Gaps: 0

US-10-631-958-2 (1-326) x US-10-631-958-4 (1-474)

Qy 80 GlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyVal 99  
Db 8 GGTATGTTCCAGCGAGGTCTGCACGGTCTGATTGGGAGACGCAGAGGAGCGCGGGTTC 67  
Qy 100 AspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIlePro 119  
Db 68 GACCAAGAACCCCGCGGCTGTGCTGGTCCCGAGTAGCTCCGGATTGGAATCATTCCTCC 127  
Qy 120 AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer 139  
Db 128 GCAGGCTCAACGACATGCGTGTGTACTCCACCGTGGGACCCAGGACGACGAACCTCG 187  
Qy 140 AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerValHisHisAsn 159  
Db 188 GCGCTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCAGTCCACCAAC 247  
Qy 160 SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIle 179  
Db 248 AGCACATCTCTCGCTACTCCGTGTCCCTGCTGGGTACGGTCTCTACGGGACATCATC 307  
Qy 180 LysAspSerGluLysIleValGlyTrpLeuGlyLeuAlaAsqTyrAspPheSerGlyLeuLys 199  
Db 308 AAGGACAGTGAGAAGAAACGGTGGTCTTGGGCTCTTGGCCAGATACGACTTTTCAGGTTTAAAG 367  
Qy 200 ThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr 219  
Db 368 ACCTTCCTCCACCACCTGCTATGAAGGACAGTGTCTCTCTCCCTGCACAACACAG 427

Qy 220 ValGlySerProArgAspArgLysProCysArgAlaGlyCysPhe 234  
Db 428 GTGGGATCTCCAAGGATAGGAAGCCCTGCGGGGCAAGATGCTTT 472  
RESULT 15  
US-10-425-115-176413  
; Sequence 176413, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: Plants  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 176413  
; LENGTH: 2657  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MMT4577\_92481C.1  
US-10-425-115-176413

Alignment Scores:  
Pred. No.: 5.6e-57 Length: 2657  
Score: 529.00 Matches: 131  
Percent Similarity: 45.97% Conservative: 57  
Best Local Similarity: 32.03% Mismatches: 107  
Query Match: 30.81% Indels: 114  
DB: 20 Gaps: 11

US-10-631-958-2 (1-326) x US-10-425-115-176413 (1-2657)

Qy 1 ProlHisLeuLeuValPheIleAsnProPheGlyLysGlyGlnGlyLysArgIle 20  
Db 556 CCAAAGAACTTAATGGTATTTGTTTCATCCATTGTGTGGAAAGGTAGAGGGTTCAGTAAC 615  
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40  
Db 616 TGGGAA---ACAGTGTATCCGTATTTCCTAGACGAAGAATAAAACAAAGTGATA--- 669  
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60  
Db 670 -----GTGACAGAGAGAGCAGGCGCATGATATGAC 699  
Qy 61 Thr-----LeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGly 77  
Db 700 ACACATCATCATTCATAGATATAGAACTGAAGGCATTTGATGGTGTGTTGTTGCAGTGGGT 759  
Qy 78 GlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArg----- 95  
Db 760 GCGCAGCGTCTATTTAATGAAATTCCTGAATGGGACTACTAGTTCTTAGGCACAAGGTTTCA 819  
Qy 95 ----- 95  
Db 820 TATCCCCCACTCCTGAGGGGTTTGGATATGTTGGACAGCTGATACTGTGAAGAACAG 879  
Qy 95 ----- 95  
Db 880 ACAAGAAATTCGGTTAACTTTAGTAAACCCACACACAGATTCCTGGGAATGCGGTGTTCTA 939  
Qy 96 -----SerAlaGlyValAspGln-----AsnHisProArgAla 106  
Db 940 GGGAAATCCAAGTAATGTGATGACCAAGAACCTCTCTTTTCACTTCATCTCCTAGCGAG 999  
Qy 107 ValLeuValProSerSer----- 112  
Db 1000 CTAGAGATACCATCAAAATTCCTACTAAAGAACCACTGATTGGAGATCAGGATAACCCAGTT 1059

```
Qy 113 -----LeuArgIleGlyIleIleProAlaGlySerThrAspCysVal 126
Db 1060 TCCTTCCCAATGACTGTTTAGGCTTGGCATATTCCTCTCGCTCAACTGATGCTATT 1119
Qy 127 CysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGly 146
Db 1120 GTTCTCAGCAACAACCTGGGGAGAGATCCTGTCACCTCTGCTCTATTATTCTTGGC 1179
Qy 147 AspSerLeuAlaMetAspValSerSerValHisIleAsnSerThr----- 161
Db 1180 AGAAGGTGTACATTGATATAGCTCAAGTCCTTAGTGGAAACTAGCCAGCAGCTGAG 1239
Qy 162 -----LeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIle 178
Db 1240 GTTTTGCCTACTGTACGCTATGCTTCATTTCAGGTTATGGCTTTTATGGAGAAGTT 1299
Qy 179 IleLysAspSerGluLysGlyArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeu 198
Db 1300 ATCAGGAGAGCGAAAGTACCGGTGGATGGGTCTGCACGTTATGACTTTTCTGGAACA 1359
Qy 199 LysThrPheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis 218
Db 1360 ATGGTCTTTCTGAAGCAGTCATATGGGCAAGGTTCCTTTCTCGATACTCCATAC 1419
Qy 219 Thr-----ValGlySerProArgAsp----- 225
Db 1420 ACTCATTCGCTTACAGCATCAGCACAGGATGATATATACTGGAGCACACCACTACAATCT 1479
Qy 226 -----ArgLysPro-----CysArgAlaGlyCysPheValCysArgGlnSer 239
Db 1480 CGTTGGAAGAAACCTCGCAAAATAATCTGCAGAACAAATTGCTTTGTATGCAAGAGGCT 1539
Qy 240 LysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp 259
Db 1540 TCAACATCTGGACAAAAATCCAGATGATGTTGCAGACAACTCCAGAACAAATATGCGAGAAC 1599
Qy 260 ValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSer 279
Db 1600 ---CAGAAAGTGGTCTGGTCCGAGGACATTTCTCAGCGTCGGTGCAGCTGTCATTCA 1656
Qy 280 CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer 299
Db 1657 TGTCGTAACGAAAGAGCCCTGATGGCTGGTAGCTGACGACACACCTTTCGGATGGTTT 1716
Qy 300 SerAspLeuIleLeuIleArgLysCys 308
Db 1717 CTTTCATCTCTTGCTCTAAGAGACTGC 1743
```

Search completed: September 6, 2005, 10:47:27  
Job time : 951.291 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 04:10:49 ; Search time 14.4126 Seconds  
(without alignments)  
2176.332 Million cell updates/sec

Title: US-10-631-958-2  
Perfect score: 1717  
Sequence: 1 PKHLLVFNPFGKGQKRI.....KCSRNFRLRHTNQDQ 326

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	486.5	28.3	549	2 T33517	hypothetical prote
2	365.5	21.3	1240	2 T05162	hypothetical prote
3	226	13.2	458	2 T38776	hypothetical prote
4	220.5	12.8	687	2 S51398	hypothetical prote
5	219.5	12.8	473	2 T19707	hypothetical prote
6	185.5	10.8	310	2 AG1665	hypothetical prote
7	181.5	10.6	624	2 S67059	hypothetical prote
8	178.5	10.4	310	2 A11293	hypothetical prote
9	131	7.6	306	2 AH1769	conserved hypotet
10	130	7.6	309	2 AH1528	conserved hypotet
11	129	7.5	303	2 F69795	conserved hypotet
12	125.5	7.3	295	2 A83894	hypothetical prote
13	125	7.3	294	2 G95120	conserved hypotet
14	123.5	7.2	309	2 AF1171	conserved hypotet
15	121	7.0	306	2 AE1394	conserved hypotet
16	121	7.0	311	2 C97990	conserved hypotet
17	116.5	6.8	364	2 F84898	hypothetical prote
18	113	6.6	295	2 D83734	hypothetical prote
19	111.5	6.5	345	2 E69678	involved in polyke
20	110	6.4	433	2 S75948	hypothetical prote
21	107	6.2	1028	2 A96719	hypothetical prote
22	105	6.1	315	2 A89978	conserved hypotet
23	105	6.1	732	2 T16422	hypothetical prote
24	102	5.9	343	2 JC7183	cathepsin Q (8C 3
25	101	5.9	309	2 H70861	hypothetical prote
26	97.5	5.7	309	2 H69995	hypothetical prote
27	95	5.5	455	1 B25493	indoleacetamide hy
28	93	5.4	650	1 JCI450	fibroblast growth
29	92.5	5.4	304	2 F72386	conserved hypotet

30	92.5	5.4	315	2 AB2166	hypothetical prote
31	91.5	5.3	299	2 E84975	hypothetical prote
32	91	5.3	338	2 G71901	hypothetical prote
33	90.5	5.3	223	2 AI2397	hypothetical prote
34	88.5	5.2	638	2 I39196	amiloride sensitiv
35	88.5	5.2	789	2 A39564	transcription repr
36	87.5	5.1	551	1 S72485	peptidylprolyl iso
37	87.5	5.1	1511	2 S60932	probable membrane
38	86	5.0	279	2 T20987	hypothetical prote
39	84.5	4.9	299	2 F85835	hypothetical prote
40	84.5	4.9	299	2 D90990	hypothetical prote
41	84.5	4.9	1041	2 T31097	chitin synthase (E
42	84	4.9	297	2 F69595	multidrug resistan
43	84	4.9	304	2 A86842	conserved hypotet
44	84	4.9	2153	2 T14893	scavenger receptor
45	83.5	4.9	328	2 S73860	6-phosphofructokin

ALIGNMENTS

RESULT 1

T33517  
hypothetical protein T10B11.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T33517  
R:Minx, P.; Kemp, K.  
submitted to the EMBL Data Library, October 1998  
A:Description: The sequence of C. elegans cosmid T10B11.  
A:Reference number: 221363  
A:Accession: T33517  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-549 <MIN>  
A:Cross-references: UNIPROT:Q9TZ11; EMBL:AF098993; PIDN:ARC67466.1; GSPDB:GN00019; CESP  
A:Experimental source: strain Bristol N2; clone T10B11  
C:Genetics:  
A:Gene: CESP:T10B11.2  
A:Map position: 1  
A:Introns: 26/1; 76/2; 109/3; 159/1; 187/3; 229/3; 353/1; 398/2; 434/1; 467/3

Query Match	28.3%	Score	486.5	DB	2	Length	549
Best Local Similarity	34.6%	Pred. No.	1.4e-34				
Matches	111	Conservative	66	Mismatches	113	Indels	31
Gaps	6						
Qy	1	PKHLLVFNPFGKGQKRIYERKVAFLFTLASITTDIIGNKFVYVYVEVITEHANOAKE	60				
Db	164	PKHLLVFNPFGKGQKRIYERKVAFLFTLASITTDIIGNKFVYVYVEVITEHANOAKE	213				
Qy	61	TYEINIDKY---DGIVCVGGDGMFSEVLHGLIGRTQSRAGVDQNHPRAVLPSSLRIGI	117				
Db	214	YIVEMPPQWSAIDGLVSVGGDGLFNELLSCALLRTQTDAGRNIDNPSSHLVTPHIREGI	273				
Qy	118	IPAGSTDCVSVTGTSDAETSALHIVVGDLSMDVSVHNSLTLRYVSLGLGVFGD	177				
Db	274	IGAGSANSIVSVTHETNDHATSAVHAIAGSCNVDVCTVHQHKLIRISANAISYGLGD	333				
Qy	178	IIKDSEKRWGLARYDFPSGLKTLFSLHHCYEGTVSFLPAQHTVGSPPDR-KPCRAGCFVC	236				
Db	334	VLRDSEYRCUPRIYQMSALRTTIRPIYKGMQVFSLSHKENVNPKQLPCLPCPCVVC	393				
Qy	237	RQSKQQLLEEOKKALYGLAEAAEDVEWQVCGKFLAINATNMSCACRRSPRGLSPAALHG	296				
Db	394	MK-----PQGNNDKYDYHWAH---EFTHVICCVIPTVTPF-----TPYGLAFTGIG	436				
Qy	297	DGSSDLILIRKCSRPNFLRFL	317				
Db	437	DGTLDLALVPRISRFRHNMQFM	457				

RESULT 2

T05162

Qy	246	BQKALYGLAEADV	EEWQ	-----	VVCGKFLAINATN	MSCACSPRG	LSPA	292
Db	543	EDFKIKYPLD	-EGIPSDWERLDN	ISN	LIGIFYTGOPY	VAADTK	-----	FFPA 590
Qy	293	AHLGDGSSDLIL	304					
Db	591	ALPSDGTMDWVI	602					



A;Cross-references: UNIPROT:Q92A05; GB:AL592022; PIDN:CAC97095.1, PID:g16414366; GSPDB:GSPDB:8  
C;Genetics:  
A;Gene:lin1865

Query Match 10.8%; Score 185.5; DB 2; Length 310;  
Best Local Similarity 23.0%; Pred. No. 1.6e-08;  
Matches 73; Conservative 46; Mismatches 114; Indels 85; Gaps 11;

Qy 2 KHLVFNPFPGKGOGKRIYERKVAPLFTLASITTDIIGNKFYNYVEVITEHANQAKET 61  
Db :  
3 KHARVIYNP-----TSGREIIKKNLADVLS-----ILEQAGYVTSAAHTTAEPGDACHA 51  
Qy 62 LYENIDKYDGI VCGDGMEFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAG 121  
Db :  
52 AEEAVRDYDLVVAAGSGDTINEVINGIAEQPYRP-----KVGIIPTG 94  
Qy 122 STDCCVCYSTGTSDAETSALHIWVGDSLAMDVSSVHHNSTLLRYSVSLLGVGFGYDIKD 181  
Db :  
95 TTDFPARAHLPVRPIVKATKIITAAGOSVANDLGKA--NDT---YFINIGGGGRUTELTYD 149  
Qy 182 --SBKKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCRQS 239  
Db :  
150 VPSRLKTMQLAYLVKG-----IEMLPS-----LKAT 177  
Qy 240 KOOLEEBOKKALYGLEAEEDVEEMVVCKFKFLAINATNMSCACRRSPRGSLSPAHLGDGS 299  
Db :  
178 KVKVEYDQ-----GVPEGEVM---FFLLGLTNSIGGFEK----IAPDAKLDDGK 219  
Qy 300 SDLILIRKCSRNFNLRFL 317  
Db :  
220 FSLIIVKKNLAEFIRLV 237

RESULT 7  
S67059  
hypothetical protein YOR17lc - yeast (Saccharomycetes cerevisiae)  
N;Alternate names: hypothetical protein O3615  
C;Species: Saccharomycetes cerevisiae  
C;Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 09-Jul-2004  
C;Accession: S67059  
R;Bordone, R.; Camasses, A.; Madania, A.; Martin, R.P.; Poch, O.; Tarasov, I.A.; Winse  
submitted to the Protein Sequence Database, July 1996  
A;Reference number: S67032  
A;Accession: S67059  
A;Molecule type: DNA  
A;Residues: 1-624 <BOR>  
A;Cross-references: UNIPROT:Q12246; EMBL:Z75078; NID:g1420415; PID:g1420417; GSPDB:GN000  
A;Experimental source: strain S288C  
C;Genetics:  
A;Gene: SGD:LCB4; MIPS:YOR17lc  
A;Cross-references: SGD:S0005697  
A;Map position: 15R

Query Match 10.6%; Score 181.5; DB 2; Length 624;  
Best Local Similarity 25.2%; Pred. No. 9.1e-08;  
Matches 62; Conservative 43; Mismatches 96; Indels 45; Gaps 7;

Qy 2 KHLVFNPFPGKGOGKRIYERKVAPLFTLASITTDIIGNKFYNYVEVITEHANQAKET 61  
Db :  
227 RSILVFIINPHGGKGTAKNFLFTKARPILVES-----GCKIEIAY----TKYARHAIDI 275  
Qy 62 LYENIDKYDGI VCGDGMEFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAG 121  
Db :  
276 AKLDJDSKYDTIACASGDGIPVEVINGLYRRPDR---VDAFNKLAVTQ-----LPCG 324  
Qy 122 STDCCVCYSTGTSDAETSALHIWVGDSLAMDV---SSVHHNSTLLRYSVSLLGYGFGYCDI 178  
Db :  
325 SGNAWSISCHWTNNPSYAALCLVKSIETFDLMCCSQPSMYNEWRPLSFLSQYGVTAES 384  
Qy 179 IKOSEKKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCRQ 238  
Db :  
385 DINTEFFRWGMGFVRFNLG-----VAPNIIOG-----KKYCEVFVKYAAK 424

Qy 239 SKOOLE 244  
||:|:  
Db 425 SKKELK 430  
||:|:  
RESULT 8  
A11293  
hypothetical protein lmol1753 [imported] - Listeria monocytogenes (strain EGD-e)  
C:Species: Listeria monocytogenes  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: A11293  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: A11293  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-310 <GLA>  
A:Cross-references: UNIPROT:Q8Y6D4; GB:NC\_003210; PIDN:CAC99831.1; PID:g16411207; GSPDB:  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmol1753  
Query Match 10.4%; Score 178.5; DB 2; Length 310;  
Best Local Similarity 22.3%; Pred. No. 6.6e-08;  
Matches 71; Conservative 49; Mismatches 113; Indels 85; Gaps 11;  
Qy 2 KHLVFNPFCKGQKRIYERKVAFLFTLASITTDIIGNKFYVYVEVITEHANQAKET 61  
||:|:  
Db 3 KHARVIYNP-----TSGREIIKKNLADVLS-----ILEQAGYVYTSAHATTAEPPDAKHA 51  
62 LYEINIDKYGIVCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPVPSLRIGIIPAG 121  
||:|:  
Db 52 AEEAVRRNFDLVAAAGDGTINEVINGTAKEYRP-----KVGIIPTG 94  
122 STDCVCYSTGTGSDAETSALHIVVGDLSLMDVSSVHHNSTLLRYSVSLGFGYGDIIKD 181  
||:|:  
Db 95 TTNDFAHALHVPDRVIKATKIIAAGQSVAMDIGKA--NET---YFINIGGGRLTETLYD 149  
182 --SEKKRWGLARDFSLGKTLFSLHCYEGTVSFPAQHTVGSPRDRKPCRCAGCFVCRQS 239  
||:|:  
Db 150 VPSRLKTMGLQAYLYLK-----IEMLPS-----LKAT 177  
240 KOOLEEOKKALYGLEAAEDVEEQVCGKFLAINATMNSCACRRSPRGLSPAHLGDGS 299  
||:|:  
Db 178 KVVEYDQ-----GVFEQV-----FFLIGLTNSIGGFKE----IAPDAKLDDGK 219  
300 SDLILRKCSRFNLRFL 317  
||:|:  
Db 220 FSLIIVKVNLAEIFRLV 237  
RESULT 9  
A11769  
conserved hypothetical protein lin2702 [imported] - Listeria innocua (strain Clip11262)  
C:Species: Listeria innocua  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: A11769  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: A11769

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-306 <GLA>  
A:Cross-references: UNIPROT:Q927T6; GB:AL592022; PIDN:CAC97928.1; PID:g16415238; GSPDB:G  
A:Experimental source: strain Clip11262  
C:Genetics:  
A:Gene: lin2702  
Query Match 7.6%; Score 131; DB 2; Length 306;  
Best Local Similarity 18.9%; Pred. No. 0.00095;  
Matches 62; Conservative 55; Mismatches 91; Indels 120; Gaps 14;  
Qy 2 KHLVFNPFCKGQKRI-----YERKVAFLFTLASITTDIIGNKFYVYVE 49  
||:|:  
Db 3 KKAMIIYNPAAGKNKFKLLPDAEKILTNADFVTLVP-STPAKPSITLI----- 51  
50 VITEHANQAKETLYEINIDKYDVCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLV 109  
||:|:  
Db 52 -----AKQAAEAGYEV-----VIAAGDGTVNEVNGLMQVEKRP----- 86  
110 PSSLRIGIIPAGSTDCVYSTGTGSDAETSALHIVVGDLSLMDVSSVHHNSTLLRYS-- 166  
||:|:  
Db 87 -----KUGILPVGITNDVARALNFANQDPLEALQIIAKOETIRVDIGKANETETFFINNAGG 142  
167 -VSLGFGYGDIIKDSEKKRWGLARDFSLGKTLFSLHCYEGTVSFPAQHTVGSPRD 225  
||:|:  
Db 143 RITEITYA-----VKESMKS KM-GRLAYLFSGL-----TVLP----- 173  
226 RKPCRACFCVCRSQKQLEEEOKKALYGLEAAEDVEEQVCGKFLAINATN-----MSCA 281  
||:|:  
Db 174 -----KLSPVNVEIIN-----BEIFKGEIL-----LFFVNKTNSVGGMETL 210  
282 CRRSPRGLSPAHLGDGSDLLILIRKCS 309  
||:|:  
Db 211 C-----PPAQLNSGMFELLILKRV 230  
RESULT 10  
A11528  
conserved hypothetical protein lin0768 [imported] - Listeria innocua (strain Clip11262)  
C:Species: Listeria innocua  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: A11528  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: A11528  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-309 <GLA>  
A:Cross-references: UNIPROT:Q92DP5; GB:AL592022; PIDN:CAC96000.1; PID:g16413219; GSPDB:G  
A:Experimental source: strain Clip11262  
C:Genetics:  
A:Gene: lin0768  
Query Match 7.6%; Score 130; DB 2; Length 309;  
Best Local Similarity 23.0%; Pred. No. 0.0012;  
Matches 46; Conservative 36; Mismatches 80; Indels 38; Gaps 8;  
Qy 5 LVFINPFGKQKRIYERKVAFLFTLASITTDIIGNKFYVYVEV-ITEHANQAKETLY 63  
||:|:  
Db 5 LLVINPSSGKEGK-TYQK-----TEEVLLKKRY--DEVEVLTEKAGDATBFAS 51  
64 EINIDKYDGVICVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPVPSLRIGIIPAGST 123  
||:|:  
Db 52 WASEQGFDAVIAMGDDGTLETINGL-----AIHEKRP-----DFGFPLGTV 94  
124 DCVCYSTGTGSDAETSALHIVVGDLSLMDVSSVHHNSTLLRYSVSLGFGYGDIIK--D 181

Db 95 NDLARSVGIPLEKXIAQALEKAKIPMDIGIGD-----QFMVNLGMAQIAVDQVS 149  
QY 182 SEKKRWGLARVDFSGLTKTF 201  
Db 150 VEQKTKFGSVAYFLEGLKAF 169

RESULT 11  
F69795  
conserved hypothetical protein yerQ - Bacillus subtilis  
C:Species: Bacillus subtilis  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C:Accession: F69795  
R:Kunet, F.; Ogasawara, N.; Mosser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertel  
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen  
isch, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Huilo, M.F.  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K  
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A69580; MUID:98044033; PMID:9384377  
A:Accession: F69795  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-303 <KUN>  
A:Cross-references: UNIPROT:031502; GB:Z99107; GB:AL009126; NID:g2632866; PIDN:CAB12492.  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yerQ

Query Match 7.5%; Score 129; DB 2; Length 303;  
Best Local Similarity 21.0%; Pred. No. 0.0014;  
Matches 70; Conservative 52; Mismatches 114; Indels 98; Gaps 15;

QY 2 KHLVFNPFPGKGKGIYERKVAFL---FTLASITTDIGNKFYVNVVEVITEHANQA 58  
Db 2 KRAIYINP-----TSGREIFKHLAQVLQKFEQAGYETSTHA-----TTCAGDA 46

QY 59 KETLEINIDKYDGIVCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLPSSLRIGII 118  
Db 47 THAAKEAALREFDLIAAGDGTINEVNGL-----APLD-NRP-----TLGVI 89

QY 119 PAGST-DCVYSTVGTSDAETSALHIVVGDLSLMDVSSVHHNSTLLRYSVSLLGYGYGD 177  
Db 90 PVGTNDFARALGIPREDILKAADTVINGVAPDIDGVQNG-----QFINIAGGERLTE 144

QY 178 IKKD--SEKKRWGLARYDFSGLTKTFLSHCHYEGTVSFPLPAQHTVGSPRDRKPCRAGCFV 235  
Db 145 LTYDVPESKLTWGLQALYVLKQME-----MLPS----- 172

QY 236 CROSKQOLEEOKKALYGLEAAEDVEWQVCGKFLAINATMNSCACRRSPRGLSPAHL 295  
Db 173 LRPTEVEYDQK--LFQGEI-----MLFLVTLTNSVGGFEK-----LAPDSSL 214

QY 296 GGGSSDLILIRKCSRFNLFRLI-----RHTNOQ 324  
Db 215 NDMGFDLMILKXANLAERFIRVATMALRGEHINDQ 248

RESULT 12  
A63894  
hypothetical protein BH1953 [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C>Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004

C:Accession: A63894  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
C:Accession: A63894  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-295 <STO>  
A:Cross-references: UNIPROT:Q9KEH4; GB:AP001513; GB:BA000004; NID:g10174345; PIDN:BA0056  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH1953

Query Match 7.3%; Score 125.5; DB 2; Length 295;  
Best Local Similarity 21.2%; Pred. No. 0.0027;  
Matches 68; Conservative 46; Mismatches 108; Indels 99; Gaps 16;

QY 6 VFINPFPGKGKGIYERKVAFLFTLASITTDIGNKFYVNVVEVITEHANQAKETLYEI 65  
Db 4 LIVNKASNGKQRTWK-----VEYELQIRNTPYLVRF-----TSGSGHATTIVKEL 51

QY 66 NIDKYDGIVCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLPSSLRIGIIPAGSTDC 125  
Db 52 LTEGVKIIIVGGDGTINEVANGIV-----NH-----RVP-----LGIIPAGS--- 89

QY 126 VCYSTVGTSDAE-----TSALHIVVGDLSLMDVSSVHHNSTLLRYSVSLLGYGYGD 177  
Db 90 -----GNDFARCLNIPMHEKALHRIF-ENKQKVDLLHLGQ---RHCLTVTGIGFGDK 139

QY 178 IIK--DSEKKRWGLARYDFSGLTKTFLSHCHYEGTVSFPLPAQHTVGSPRDRKPCRAGCF 234  
Db 140 IAKTVNEAIYNWFN--QFGFGLSYVLS-----MLEVLDKDYRP----- 176

QY 235 VCROSKQOLEEOKKALYGLEAAEDVEWQVCGKFLAINATMNSCACRRSPRGLSPAHH 294  
Db 177 ----TNIQITVDGKELFFS-----GVWLVA-----VANSFNYGGGIR-----ICPEAS 215

QY 295 LGDSSDLILIRKCSRFNFLR 315  
Db 216 YDDGLNLICVVGMSKWLRLR 236

RESULT 13  
G95120  
conserved hypothetical protein TIGR00147 SP1045 [imported] - Streptococcus pneumoniae (s  
C:Species: Streptococcus pneumoniae  
C>Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004  
C:Accession: G95120  
R:Tetzelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,  
nson, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
A:Title: Complete genome sequence of a virulent isolate of Streptococcus pneumoniae.  
A:Reference number: A95000; MUID:21357209; PMID:11463916  
C:Accession: G95120  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-294 <KUR>  
A:Cross-references: UNIPROT:Q9TQZ6; GB:AB005672; PIDN:AAK75160.1; PID:g14972520; GSPDB:G  
A:Experimental source: strain TIGR4  
C:Genetics:  
A:Gene: SP1045

Query Match 7.3%; Score 125; DB 2; Length 294;  
Best Local Similarity 23.4%; Pred. No. 0.003;  
Matches 49; Conservative 37; Mismatches 85; Indels 38; Gaps 8;

QY 2 KHLVFNPFPGKGKGIYERKVAFLFTLASITTDIGNKF--VNVYVEV-ITEHANQA 58  
Db 2 KXAVIINPISG-GEKALDYKEK-----LENKAKYFEFVETKITEKALDA 46



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 6, 2005, 07:01:07 ; Search time 3134.86 Seconds  
(without alignments)  
3958.373 Million cell updates/sec

Title: US-10-631-958-2  
Perfect score: 1717  
Sequence: 1 PKHLLVFNPFGCKGQGRRI.....KCSRNFRLFLRHTNQDQ 326

Scoring table:

BLSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US10631958/runat\_02092005\_165816\_4562/app.query.fasta\_1.1941  
-DB=BST\_QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosom62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10631958 @CGN 1 1 9235 @runat\_02092005\_165816\_4562 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsa1:\*  
9: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Query	ID	Description
1	1500.5	87.4	1824	3 AK042077	Mus muscu
2	1500.5	87.4	4248	3 AK052269	Mus muscu
3	1455.5	84.8	1063	4 BM479389	AGENCOURT
4	1350.5	78.7	1059	5 BQ054406	AGENCOURT
5	1283	74.7	797	7 CK000755	AGENCOURT
6	1267.5	73.8	1047	5 BQ057191	AGENCOURT
7	1176	68.5	758	6 CB246749	UI-M-F10-
8	1073.5	62.5	1078	5 BQ063738	AGENCOURT
9	1041	60.6	820	6 CD655311	AGENCOURT

10	1012	58.9	581	5	BP224560	BP224560
11	980.5	57.1	584	5	BP310011	BP310011
12	965.5	56.2	584	5	BP309990	BP309990
13	951.5	55.4	573	7	CF138275	UI-HF-BN0
14	942	54.9	541	7	CN296311	170005321
15	925	53.9	725	5	BU214294	603756937
16	917	53.4	550	7	CF138634	UI-HF-BN0
17	916	53.3	576	6	CB608292	AMGNNUC:N
18	905.5	52.6	779	5	BQ265377	603374309
19	898	52.3	713	7	CK000525	AGENCOURT
20	892	52.0	661	7	CN296312	170005321
21	873	50.8	674	1	AJ739034	AJ739034
22	861.5	50.2	553	6	CA578972	K0731A08-
23	831	48.4	545	6	CB611947	AMGNNUC:N
24	828	48.2	488	5	BU430459	UI-HF-BN0
25	811	47.2	474	2	AW503999	UI-HF-BN0
26	809	47.1	902	5	BU373554	603811294
27	795	46.3	653	5	BX952302	DKF20781L
28	761	44.3	759	5	BU290807	604162346
29	749	43.6	759	5	BU264162	603814452
30	741	43.2	732	7	CF135528	UI-HF-BN0
31	705.5	41.1	735	7	CF535441	UI-M-GH0-
32	695.5	40.5	467	4	BM484519	538413 MA
33	694.5	40.4	638	1	AJ739022	AJ739022
34	688.5	40.1	452	6	CA579952	K0744G05-
35	640	37.3	376	5	BU430866	UI-HF-BN0
36	637.5	37.1	819	7	CN235868	WLB086C10
37	627	36.5	567	1	AJ739023	AJ739023
38	591	34.4	359	5	BU428926	UI-HF-BN0
39	588	34.2	763	7	CK364327	AGENCOURT
40	587.5	34.2	725	5	BU225411	603399474
41	570	33.2	607	6	CB581157	AMGNNUC:N
42	570	33.2	656	4	BJ673686	BJ673686
43	555	32.3	660	4	BJ620045	BJ620045
44	508.5	29.6	329	1	AA355581	EST64087
45	500.5	29.1	310	6	CD631566	56095954H

#### ALIGNMENTS

RESULT 1	AK042077	1824 bp	mRNA	linear	HTC 03-APR-2004
LOCUS	AK042077	Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched library, clone:A630056D11 product:DA59H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS)			
DEFINITION	AK042077	(FRAGMENT) homolog [Homo sapiens], full insert sequence.			
ACCESSION	AK042077.1	GI:26334912			
VERSION	AK042077.1	GI:26334912			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Carninci,P. and Hayashizaki,Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	1	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.			
AUTHORS	2	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
TITLE	Genome Res. 10 (10), 1617-1630 (2000)				
JOURNAL	20499374				
MEDLINE	11042159				
REFERENCE	3	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,			

Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A., and Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multipillar sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

## TITLE

JOURNAL  
MEDLINE  
PUBMED

## REFERENCES

## AUTHORS

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

## TITLE

Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)

## REFERENCE

## AUTHORS

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

## TITLE

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

## JOURNAL

## REFERENCE

## AUTHORS

6 (bases 1 to 1824)  
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saichou,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

## TITLE

## JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216]

## COMMENT

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.  
Please visit our web site for further details.  
URL:http://genome.gsc.riken.jp/  
URL:http://fantom.gsc.riken.jp/.

## FEATURES

## source

location/Qualifiers  
1..1824  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="PANTOM\_DB:A630056D11"  
/db\_xref="taxon:10090"  
/clone="A630056D11"  
/tissue\_type="thymus"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="3 days neonate"  
85..1680  
/note="unnamed protein product; DA59H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEIN) (FRAGMENT) homolog [Homo sapiens] (SPTF|O9UE5, evidence: FASTA, 86.5%ID, 100%length, match=945) putative"  
/codon\_start=1  
/protein\_id="BAC31157.1"  
/db\_xref="GI:26334913"  
/translation="MGAMGAAEPLHSLVWVKRRRCVSLSEPARALLRWRRSPPEPQPSA

## CDS

ORIGIN  
Alignment Scores:  
Pred. No.: 1,04e-154 Length: 1824  
Score: 1500.50 Matches: 282  
Percent Similarity: 92.64% Conservative: 20  
Best Local Similarity: 86.50% Mismatches: 13  
Query Match: 87.39% Indels: 11  
DB: 3 Gaps: 1  
US-10-631-958-2 (1-326) x AK042077 (1-1824)  
Qy 1 ProlyserHisLeuLeuValPheIleAsnProPheGlyGlyGlyGlnGlyLysArgIle 20  
Db 472 CCGAAGCAGCTTGGTGGTATTTCATCAACCCCTTCGAGGGAAGGTACGGGCAACGCCATC 531  
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40  
Db 532 TATGAAAAACAGTCGGCGCTCTGTTTACCTTGGCTCCATCTACGAGATCATC-- 588  
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60  
Db 589 -----ATTACAGAGCATGCCAACCAACGCAAGGAG 618  
Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80  
Db 619 ACTTTATACAGAGATCAACACAGACAGCTATGATGGCATGTGTGCTAGGTGGGAGCGGC 678  
Qy 81 MetPheSerGluValLeuHisGlyIleGlyArgThrGlnArgSerAlaGlyValAsp 100  
Db 679 ATGTTACGAGGTTGCTGCGTGGGTGTTGGGAGGACGAGCAGCGCTGGTATCGAC 738  
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120  
Db 739 CCCAATCACCCCGAGCGGTGCTGGTCCCGCATCCCTCAGGATCGGCATCATACCCGCA 798  
Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140  
Db 799 GGGTCCACAGATTGTGTGTACTCAACAGTGGGCACAAACGACGAGACAGATCGGCT 858  
Qy 141 LeuHisIleValValGlyAspSerSerLeuAlaMetAspValSerSerValHisHisAsnSer 160  
Db 859 TTGCACATCATTTATGGGACTCATTCGCAATAGACGTGCTCTGTGCTACTACCAATAC 918  
Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180  
Db 919 ACGTGTGCGGTACTCGGTTTCTCTGCTGGGTACCGTTTCTACGGGGAGTTAATCAAG 978  
Qy 181 AspSerGluLysLysArgTyrLeuLeuAlaArgTyrAspPheSerGlyLeuLysThr 200  
Db 979 GACAGTGAAGAAGAACGGTGGATGGGCTCTGTCGGGTATGATTTCTCAGGGTTGAAGACC 1038  
Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220  
Db 1039 TTTCTCTCTCATCATGACTATGAAGGACACTGCTCTTCTCCAGCAGACAGCACACGGTG 1098  
Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240  
Db 1099 GGATCTCCACGGGCAATAAACCCCTGCGGGCTTCGGTGGTGGTGGTGGTGGTGGTGGTGGT 1158  
Qy 241 GlnGlnLeuGluGluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260  
Db 1159 CACAGCTGGGAAGAAGAGAGAGAACGCCCTGTATGGCTTGGAGAGACGCCGAGGAATG 1218

PGADARSVLVSEIIAIVEEKDDCEKSHSSGRWHKVENPFAFVHRVKKVRHHRWKWARV  
TFSADEQLCHLWQLTLLGLESITSRPKLLVFINPFGCGQCKRIYETKIVAPLPLFL  
ATSTTEIITTEHANOAKETIYEINTSDYGVICVGGDMFSEVLGHVIGRTQOSAGID  
PNHPRALVPSTLRIIGIPAGSTDCVCYKVTGNDAAETSAHIIIGDSLADYSSVHV  
HNTLLRSVSLGKGYPGDLIKQSEKRWGNGLVRYDFSGLTKFLSHQYEGTUSFLPA  
QHTVGSPRNDKCRAGFCVRQSKQLEBEKALYGLNEAEEMEEQVTCGKPLAIN  
ATNMSCAPRSPGLSPFAHLGSGSDLILIRKSRNFRLRLRIHTNNDQDFDTFV  
EYVRVKKFHTSKHVEDENDNSKEQBKFQKICKDRPSCSASRSSMNCDEGVMS  
PAIEVRVHCQLVRLFARGIEES"

QY 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280  
 |||||  
 Db 1219 GAAGAGTGGCAAGTGCATGTGGAGATTCCTGGCCATCAATGCCACCAATGCTCTGT 1278

QY 281 AlaCysArgSerProArgGlyLeuSerProAlaIleHisLeuGlyAspGlySerSer 300  
 |||||  
 Db 1279 GCTTGTCTCGAGCCCTGGGGCCCTGTGCCATTTGCCCATCTGGGAGATGGTCTTCT 1338

QY 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320  
 |||||  
 Db 1339 GACCTCATCTTATCCGGAAGTCTCCAGGTTCAATCTCTGAGATTCTCATCCGGCAC 1398

QY 321 ThrAsnGlnAspGln 326  
 |||||  
 Db 1399 ACGAACACGAGGACGAC 1416

RESULT 2  
 AK052269  
 LOCUS  
 DEFINITION Mus musculus 13 days embryo heart cDNA, RIKEN full-length enriched library, clone:D330016D08 product:DA59H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS) (FRAGMENT) homolog [Homo sapiens], full insert sequence.  
 AK052269  
 ACCESSION AK052269.1 GI:26342491  
 VERSION  
 KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

REFERENCE  
 AUTHORS Carninci, P. and Hayashizaki, Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636

REFERENCE  
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20493374  
 PUBMED 11042159

REFERENCE  
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Ozawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE 20530913  
 PUBMED 11076861

REFERENCE  
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the Functional annotation of a full-length mouse cDNA collection  
 JOURNAL Nature 409, 685-690 (2001)  
 MEDLINE 11076861  
 PUBMED  
 TITLE The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
 ANALYSIS Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 JOURNAL Nature 420, 563-573 (2002)  
 MEDLINE 120530913  
 PUBMED 120530913

REFERENCE  
 AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,

Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahiza, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

# TITLE

## JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

# COMMENT

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Please visit our web site for further details.  
 URL: http://genome.gsc.riken.jp/  
 URL: http://fantom.gsc.riken.jp/

# FEATURES

## Location/Qualifiers

1..4248  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM.DB:D330016D08"  
 /db\_xref="taxon:10090"  
 /clone="D330016D08"  
 /tissue\_type="heart"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stages="13 days embryo"  
 260..1855

## CDS

/note="unnamed protein product; DA59H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS) (FRAGMENT) homolog [Homo sapiens] (SPTR|09UG85, evidence: FASTY, 86.5%ID, 100%length, match=945) putative"  
 /codon\_start=1  
 /protein\_id="BAC34908.1"  
 /db\_xref="GI:26342492"  
 /translation="MGAGAAAPLHSLVWKRRCVSLSEPARALLRWRSPSPQPSA PGADARSVLSEIIAVEEKDDCEKHSRWHKMPFAFVHRVYRHRHWKWAR ASITTEILITEHANOAKETLYEINTDSYDGVCGDGMFSEVLHGIVRGTQOASGID PNHPRVLVPTSLIGIIPAGSTDCVSTGTNDASTALHIIIGSLAIDVSVHY HNTLLRYSVLGIGYFDGLIKDSEKRWMLVRYDFSLKTFILSHQYEGTSLFLPA QHTVGSPRDNKPCRGAGFCVRSQKQLEEEKALYGLENAEEMEQVTCGKFLAIN ATNMSCAPRSPGGLSPFAHGDGSSDLILRKCSRFNLFRLIRHTNOEDQDFTFV EYVRVKPFHTSKHVEYEDNDSDKEQKFGKICKDRPSCCTCSASRSSWNCDEVMHS PAIEVRVHCQLVRLFARGIBES"

## ORIGIN

Alignment Scores:  
 Pred. No.: 3, 7e-154 Length: 4248  
 Score: 1500.50 Matches: 282  
 Percent Similarity: 92.64% Conservative: 20  
 Best Local Similarity: 86.50% Mismatches: 13  
 Query Match: 87.39% Indels: 11  
 DB: 3 Gaps: 1  
 US-10-631-958-2 (1-326) x AK052269 (1-4248)

QY 1 ProlysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgile 20  
 |||||  
 Db 647 CCGAGCACTTGTCTGGTATTCATCAACCTTTCGAGGAGAAAGGTCAGGCAAGCGCATC 706  
 |||||  
 QY 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40  
 |||||  
 Db 707 TATGAAAAACAGTGGCCCTCTGTTTACCTTGGCTTCCATCCTACCGAGATCATC--- 763





```

Db      448  CGGTGGTGGGCTTGGCCAGATACGACTTTTCAAGTTTAAAGACCTTCTCTCCACCAC 507
Qy      206  CysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAep 225
Db      508  TGTATGAAGGGACAGTGTCTTCTCTCCCTGCACACACACACGCTGGGATCTCCAAGGAT 567
Qy      226  ArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnLeuGlu 245
Db      568  AGAAGCCCTGCCGGCAGATGCTTTGTTGAGCAAGCAAGCAGCAGCTGGAGGAG 627
Qy      246  GluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlnVal 265
Db      628  GAGCAGAAGAAGCACTGTATGGTTTGAAGCTGCCGAGGAGCTGGAGGAGTGGCAAGTC 687
Qy      266  ValCysGlyLysPheLeuAlaLeuAlaThrAsnMetSerCysAlaCysArgArgSer 285
Db      688  GTCTGTGGGAAGTTTCTGCNCATCATGCCACAAACATGCTCCTGTGTGCTGCCGGAGC 747
Qy      286  ProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuLeuLeu 305
Db      748  CCCANGGGGCTCTCNCGCCCTGCCACCTTGNAGAGCGGTCTTCTGACCTCATCTCATC 807
Qy      306  ArgLysCysSerArgPheAsnPheLeuArgPheLeuLeuArgHisThrAsn-GlnGlnAs 325
Db      808  CGGAATGCTCCAGGNYCAATTTCTGAGATTTCTCATCAGGCACACCAACCCAGCAGGA 867
Qy      325  p 325
Db      868  C 868

```

## RESULT 4

BQ054406

LOCUS

DEFINITION

BQ054406 1059 bp mRNA linear EST 29-MAR-2002

5', mRNA sequence.

ACCESSION BQ054406

VERSION BQ054406.1

KEYWORDS GI:19813746

SOURCE EST.

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1059)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaabs-r@mail.nih.gov

Tissue Procurement: Iou Staudt

cDNA Library Preparation: Rubin Laboratory

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Clone Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCM2040 row: a column: 13

High quality sequence stop: 601.

Location/Qualifiers

1..1059

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5803668"

/tissue\_type="lymphoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_99"

/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:

ECORI; cDNA made by oligo-dt priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:

GGCAGCAG(G). Size-selected &gt;500bp for average insert size

1.8kb. Library constructed by Ling Hong in the laboratory

of Gerald M. Rubin (University of California, Berkeley)

## FEATURES

source

using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
II RT (Life Technologies). Note: this is a NIH\_MGC  
Library."

## ORIGIN

```

Alignment Scores:
Pred. No.: 1,9e-138 Length: 1059
Score: 1350.50 Matches: 275
Percent Similarity: 89.49% Conservative: 6
Best Local Similarity: 87.58% Mismatches: 14
Query Match: 78.65% Indels: 19
DB: 5 Gaps: 5

US-10-631-958-2 (1-326) x BQ054406 (1-1059)

Qy      4  LeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArg 23
Db      15  TTACTGTGTTATTTATCAACCGCTTTGGAGGAAAGGACAAAGGCAAGCGCATATATGAAAGA 74
Qy      24  LysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGlyAsnLysPhe 43
Db      75  AAAGTGGCACCACATGTTTCCACCTTAGCCTCCATCACCACCTGACATCATC----- 122
Qy      44  TyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGluThrLeuTyr 63
Db      123  -----GTTACTGAAACATGCTAATCAGGCGCAAGGAGACTCTGTAT 161
Qy      64  GluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPheSer 83
Db      162  GAGATTACATAGACAAATACGACGGCATCTGCTGTGCGCGGAGATGTTATGTTTCAGC 221
Qy      84  GluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHis 103
Db      222  GAGGTGCTGCACGGTGTGATTGGGAGGACGAGAGGCGCGGGGTGCACAGAACCCAC 281
Qy      104  ProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySerThr 123
Db      282  CCCCGGGCTGTGTTGTTCCCGATAGCTCCGGATTGGAATCATTTCCCGCAGGGTCAACG 341
Qy      124  AspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIle 143
Db      342  GACTCGGTGTGTTACTCCACCGTGGCCACGACGACGAGCAAAACCTCGCGCTGCATATC 401
Qy      144  ValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeuLeu 163
Db      402  GTTGTGGGAGCTCGCTGGCCATGATGTCTCCTCAGTCCACCAACACAGCACACTCCTT 461
Qy      164  ArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGlu 183
Db      462  CGCTACTCCGTGCTCCCTGCTGGGCTTACGGGTCCTACGGGGACATCATCAAGGACAGT 521
Qy      184  LysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSer 203
Db      522  AAGAAACGGTGTGGTGGTCTTCCAGATACACATTTTCAGGTTTAAAGACCTTCTCTCC 581
Qy      204  HisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerPro 223
Db      582  CACCACCTCTATGAACGACAGTGTCTTCTCCTCCGCAACACACACGCTGGGATCTCCA 641
Qy      224  ArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnLeu 243
Db      642  AGGGATAGGAAGCCCTTCCCGGCGAGGATGCTTTGTTGAGGCAAGCAAGCAGCAGCTG 701
Qy      244  GluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlnTyr 263
Db      702  GAGGAGGACAGAGAAGACATCTGTATGGTGGTGGAACTGCGGAGGACGCTGGAGGAATG 761
Qy      264  GlnValValCys-GlyLysPheLeuAlaIleAsnAlaThr-AsnMetSerCysAla----C 282
Db      762  CAAGTCGTCTGTGGGAAAGTTCTCTGGCCATTAAATGCCCAAAACATGTCCTGTGCTGT 821
Qy      282  ysArgArgSerProArgLysLeuSerProAlaAla-----HisLeuGly---AspGlyS 299
Db      :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::

```



cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
<http://image.llnl.gov>  
Plate: LCM2062 row: 1 column: 15  
High quality sequence stop: 535.

**FEATURES**

## ORIGIN

Alignment Scores:			
Pred. No.:	2,91e-129	Length:	1047
Score:	1267.50	Matches:	273
Percent Similarity:	88.54%	Conservative:	13
Best Local Similarity:	84.52%	Mismatches:	13
Query Match:	73.82%	Indels:	24
DB:	5	Gaps:	3
US-10-631-958-2 (1-326) x B0057191 (1-1047)			

primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pVX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CAGCACGAC. This library was created for the University Iowa Brain Anatomy Project (BNAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

## ORIGIN

Alignment Scores:  
Pred. No.: 2,44e-119 Length: 758  
Score: 1176.00 Matches: 218  
Percent Similarity: 95.5% Conservative: 18  
Best Local Similarity: 88.2% Mismatches: 11  
Query Match: 68.4% Indels: 0  
DB: 6 Gaps: 0

US-10-631-958-2 (1-326) x CB246749 (1-758)

```
Qy 80 GlyMetPheSerGluValLeuHisGlyLeuLeuGlyArgThrGlnArgSerAlaGlyVal 99
Db 1 GGCATGTTTCAGCGAGGTCTGCATGGGTGATTTGGGAGGAGCGACGAGCGCTGGTATC 60

Qy 100 AspGlnAenHisProArgAlaValLeuValProSerSerLeuArgGlyLeuPro 119
Db 61 GACCCCAATACCCCGGAGCGGTGCTGTGCCAGTACCCCTCAGATCGGCATCATACCC 120

Qy 120 AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer 139
Db 121 GCAGGTCACAGATTGTGTGTACTACACAGTGGGCACAAACGACGACGACATCG 180

Qy 140 AlaLeuHisLeuValGlyAspSerLeuAlaMetAspValSerSerValHisAen 159
Db 181 GCTTTGCACATCATTTATTTGGGACTCACTGGCAATAGACGTCTCTCTGTGCACCTACC 240

Qy 160 SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrPheTyrGlyAspLeuLeu 179
Db 241 AACACGCTGCTGCGGTACTCGGTCTCTCTGCTGGGCTACGGTTCTCTCGGGAGCTTAATC 300

Qy 180 LysAspSerGluLysLeuArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys 199
Db 301 AAGGACAGTGAAGAAACAGGTGGATGGCCCTCGTCCGCTATGATTTCTCAGGGTTGAAG 360

Qy 200 ThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr 219
Db 361 ACCTTTCTCTCATCAGTACTATGAGGGACACTGTCTCTCTCCACGACGACACG 420

Qy 220 ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer 239
Db 421 GTGGGATCTCCACGGGAACAATAACCCCTCGCGGCTGGGTCTCTGTGTGAGGAGGAGC 480

Qy 240 LysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp 259
Db 481 AAGCAACAGCTGGAAGAGAGAGAGAGAAAGCCCTGTATGGCTGGAGAAACGCCGAGAA 540

Qy 260 ValGluGluTrpGlnValValCysGlyLysPheLeuAlaLeuAenAlaThrAsnMetSer 279
Db 541 ATGGAAGAGTGGCAAGTGACATGTGGGAAGTTCTCTGGCCATCAATGCCACCAACATGTCC 600

Qy 280 CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer 299
Db 601 TGTGCTGTCTCGAGGCCCTCGGGGCGCTGTCCCATTTGCCATTCCTGGAGATGGGTCT 660

Qy 300 SerAspLeuLeuLeuArgLysCysSerArgPheAenPheLeuArgPheLeuLeuArg 319
Db 661 TCTGACCTCATCTTATCCGGAAGTCTCCAGGTTCCANCTTCTCTGAGATTCTCTCATCCG 720

Qy 320 HisThrAenGlnAspGln 326
Db 721 CACACGAACACGAGGAGGACG 741
```

## RESULT 8

BQ063738

LOCUS

DEFINITION

AGENCOURT\_6873251 NIH\_MGC\_99 Homo sapiens cDNA clone IMAGE:5925382

5', mRNA sequence.

ACCESSION

BQ063738

VERSION

BQ063738.1 GI:19891754

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (Bases 1 to 1078)

AUTHORS

NIH-MGC http://mgc.nci.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-x@mail.nih.gov

Tissue Procurement: Lou Staudt

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LICM2096 row: p column: 23

High quality sequence stop: 640.

FEATURES

Location/Qualifiers

1..1078

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5925382"

/tissue\_type="lymphoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_99"

/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:

EcoRI; cDNA made by oligo-dT priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:

GGACGAG(G). Size-selected &gt;500bp for average insert size

1.8kb. Library constructed by Ling Hong in the laboratory

of Gerald M. Rubin (University of California, Berkeley)

using ZAP-cDNA synthesis kit (Stratagene) and Superscript

II RT (Life Technologies). Note: this is a NIH\_MGC

Library."

## ORIGIN

Alignment Scores:

Pred. No.: 9,35e-108 Length: 1078

Score: 1073.50 Matches: 219

Percent Similarity: 87.16% Conservative: 5

Best Local Similarity: 85.21% Mismatches: 15

Query Match: 62.52% Indels: 18

DB: 5 Gaps: 5

US-10-631-958-2 (1-326) x BQ063738 (1-1078)

Qy 1 ProLysHisLeuLeuValPheIleAsnProPheGlyLysGlyGlnGlyLysArgIle 20

Db 121 CCAAAGCATTTTACTGGTATTTATCAACCCGTTTGGAGGAAAGGACAGGCGGAGTA 180

Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40

Db 181 TATGAAGAAGGAGGCGGACCACTGTTCCCTTAGCTCCATCACCATGACATCATC--- 237

Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60

Db 238 -----GTTACTGAACATGCTTAATCAGGCCAAGGAG 267

Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyLysValGlyValGlyAspGly 80

Db 268 ACTCTGTATGAGATTAAACATAGACAAATACGCGCATCGTCTGTCTGCGCGGAGATGCT 327

```

QY      81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db      328 ATGTTCCAGGAGGTCTGTCACCGGTCTGATTGGAGGACGACAGAGCGCGCGGTTCGAC 387

QY      101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAla 120
Db      388 CAGAACCCACCCCGGGCTGTCTGGTCCCGACAGCTCCGATGGATCATCTCCGCA 447

QY      121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db      448 GGGTCAACCGGACTGGCGTGTGTACTCCACCGTGGGCACCGAGCGCAGAAACCTCGGCG 507

QY      141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisAsnSer 160
Db      508 CTGCATATCGTTGTGGGACCTGCTGGCCATGGATGTCTCAGTCCACACACAGC 567

QY      161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db      568 ACACTCTTCGTACTCTCCGTCTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAG 627

QY      181 AspSerGluLysGlyArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db      628 CACAGTGAAGAAGACGGTGGTGGTCTTGCAGATACGACTTTTCAGGTTAAAGACC 687

QY      201 PheLeuSerHisGlyCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db      688 TTCTCTCTCCACCATCTGTATGAAGACAGCGGTCTTCTCTCCCTGCGACACACACAGCGGT 747

QY      221 GlySer----ProArgAsp-ArgLysProCysArgAla---GlyCys-----PheValCy 236
Db      748 GGGGATCTCCAGGGAATPAGAGAGCCCTGCGCGGCGCAGGAGTGCCTTTGGTTGCCAG 807

QY      236 sArgGlnSerLysGlnGlnLeuGluGlu-----GlnLysLysAla 250
Db      808 GCCAAGCAAGCAGCCAGCTTTGGAAGGCGAGGAGGCCCAAGAAAGGCC 856

```

## RESULT 9

```

CD655311          820 bp      mRNA      linear      EST 18-JUN-2003
LOCUS              AGENCOURT 14552675 NIA Human H1 Embryonic Stem Cell cDNA Library
DEFINITION          (Long) Homo sapiens cDNA clone IMAGE:30426593 5', mRNA sequence.
ACCESSION           CD655311
VERSION             CD655311.1  GI:31895467
KEYWORDS            EST.
SOURCE              Homo sapiens (human)
ORGANISM            Homo sapiens
                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE            1 (bases 1 to 820)
AUTHORS              NIH-MGC http://mgi.nci.nih.gov/.
TITLE                National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL              Unpublished (1999)
COMMENT              Contact: Daniela S. Gerhard, Ph.D.
                    Office of Cancer Genomics
                    National Cancer Institute / NIH
                    Bldg. 31 Rm10A07 Bethesda, MD 20892
                    Email: gcgabs-r@mail.nih.gov
                    Tissue Procurement: Irene Ginis and Mahendra Rao, NIA
                    cDNA Library Preparation: Yulan Piao and Minoru Ko
                    cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                    DNA Sequencing by: Agencourt Bioscience Corporation
                    Clone distribution: MGC c lone distribution information
                    can be found through the I.M.A.G.E. Consortium/LLNL at:
                    http://image.llnl.gov
                    Plate: NDAM512 row: k column: 18
                    High quality sequence stop: 673.
                    Location/Qualifiers
                    1..820
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:30426593"

```

## FEATURES

source

```

/tissue_type="Embryonic Stem cells"
/cell_line="WA01"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"
/Note="Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01/H1 cell line. Undifferentiated human ES cell line WA01/H1 was obtained from WiCell Research Institute, Inc., Madison, WI, cultured according to their instructions, on MEF feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UTF, TERT, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLK1, SSEA-1, TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days after plating), the ES cells from 4 x 6cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with Trizol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID:11544199]) Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-TGACTAGTCTAGATCGGAGCGCGCCCTTTTCTTTT-3'] from 3.4g of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker L1-SalI, purified by phenol/chloroform extraction, and separated from free linkers by Centricion-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Tag polymerase (Takara) with a primer Sal4-S for 25 cycles. The products were purified by phenol/chloroform extraction and Centricion-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The average insert size is about 3.6kb."

```

## ORIGIN

## Alignment Scores:

```

Pred. No.:      2,47e-104      Length:      820
Score:          1041.00      Matches:      197
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.49%      Mismatches: 0
Query Match:     60.63%      Indels: 0
DB:              6              Gaps: 0

```

US-10-631-958-2 (1-326) x CD655311 (1-820)

```

QY      129 SerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSer 148
Db      12  TCCACCGTGGGCACGACGACGCAAAACCTCGCGGCTGCATATCGTTGTGGGACTCG 71

QY      149 LeuAlaMetAspValSerSerValHisAsnSerThrLeuLeuArgTyrSerValSer 168
Db      72  CTGGCCATGGATGTCTCAGTCCACCAACAGCACACTCTCTCGTACTCTCGGTGCC 131

QY      169 LeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGlyLysGlyArgTrpLeu 188
Db      132  CTGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAAACCGGTGGTTG 191

QY      189 GlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisGlyCysTyrGlu 208
Db      192  GGTCTTGCCAGATACGACTTTTTCAGGTTAGAGACCTTCTCTCCACCATCTGCTATGAA 251

QY      209 GlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysPro 228
Db      252  GGGACAGTGTCTCTCTCCCTCCCTGCACACACAGGTGGGATCTCCNAGGGATAGGAGGCC 311

QY      229 CysArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLeuGluGluGlnLys 248
Db      312  TGCCGGCGCAGCATGCTTTGTTTTCAGGCAAGCAAGCAGCAGCTGGAGGAGGAGCAGAAG 371

QY      249 LysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGluTrpGlnValValCysGly 268

```



372 AAAGCAGCTGATGTTTGGAGAGCTGCGAGAGACGCTGGAGGAGTGGCAAGTGGTGTGTGGG 431  
 269 LysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGly 288  
 432 AAGTTTCTGGCCATCAATGCCACAAACATGTCCTGTGCTTGCGCCGAGGCCCGAGGGGC 491  
 289 LeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIleLeuAlaArgLysCys 308  
 492 CTCTCCCGGCTGCCCATCTGGGAGACGGGTCTTCTGACCTCATCTCATCCGGAATGC 551  
 309 SerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsnGlnAspGln 326  
 552 TCCAGGTTCATATTTCTGAGATTCTCATCAGGCACCAACACGAGGACCAG 605  
  
 RESULT 10  
 BP224560 581 bp mRNA linear EST 15-SEP-2004  
 LOCUS BP224560 Sugano cDNA library, lymphocyte Daudi Homo sapiens cDNA  
 DEFINITION clone DAT01067, mRNA sequence.  
 ACCESSION BP224560  
 VERSION BP224560.1 GI:52097465  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 581)  
 AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,  
 Mizushima-Yugano, J., Nakai, K. and Sugano, S.  
 TITLE Sequence comparison of human and mouse genes reveals a homologous  
 block structure in the promoter regions  
 JOURNAL Genome Res. 14 (9) 1711-1718 (2004)  
 COMMENT Contact: Yutaka Suzuki  
 Department of Virology  
 Institute of Medical Science, University of Tokyo  
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
 Email: yuzuki@ims.u-tokyo.ac.jp.  
 FEATURES  
 source  
 1..581  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="DAT01067"  
 /cell\_type="lymphocyte"  
 /cell\_lines="Daudi"  
 /clone\_lib="Sugano cDNA library, lymphocyte Daudi"  
 /note="Burkitt's lymphoma"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,4e-101 Length: 581  
 Score: 1012.00 Matches: 192  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.48% Mismatches: 0  
 Query Match: 58.94% Indels: 0  
 DB: 5 Gaps: 0  
 US-10-631-958-2 (1-326) x BP224560 (1-581)  
 66 AsnIleAspLysThrAspGlyIleValCysValGlyGlyAspGlyMetPheSerGluVal 85  
 3 AACATAGCAANTACGACGGCATCGTCTGTCTCGCGGAGATGTTGTTTCAGCGAGGTG 62  
 86 LeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArg 105  
 63 CTGCACGGTCTGATTGGGAGGACGCGAGGAGCGCGGGGTGACACGAGCACCCCGG 122  
 106 AlaValLeuValProSerSerIleuArgIleGlyIleLeuProAlaGlySerThrAspCys 125  
 123 GCTGTGCTGGTGGTCCCGAGTAGCTCCGATTGGAATTCATTCGCGAGGGGTCAACGACTGC 182  
 126 ValCysThrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisGlyIleVal 145



QY 30 ThrLeuAlaSerIleThrThrAspIleIleGlyAsnLysPheTyrValAsnTyrValGlu 49  
Db 61 ACCTTAGCCTCCATCACCCTGACATCATC----- 90  
QY 50 ValIleThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLys 69  
Db 91 ---GTTACTGAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAA 147  
QY 70 TyrAspGlyIleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeu 89  
Db 148 TACGACGGCATGCTGTCTCGCGGAGATGGTATGTTTCAGCGAGGTGCTGCACGGTCTG 207  
QY 90 IleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuVal 109  
Db 208 ATTGGAGAGCGCACAGGAGCGCGGGGTGACAGAACCTCCCGGCTGTGTCTGGTCTG 267  
QY 110 ProSerSerLeuArgIleGlyIleIleProAlaGlySerThrAspCysValCysTyrSer 129  
Db 268 CCCAGTAGCCTCCGATTCGATCATTCCTCCGAGGGTCAACGGACTGCGTGTACTCC 327  
QY 130 ThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeu 149  
Db 328 ACCGTGGGACCAAGCAGCAGAACCTCGCGCTGCATATCGTTGTTGGGACTCGCTG 387  
QY 150 AlaMetAspValSerSerValHisAsnSerThrLeuLeuArgTyrSerValSerLeu 169  
Db 388 GCATGGAGTGTCTCAGTCACCAACAGACACTCCTTCGCTACTCCGTGTGCCTG 447  
QY 170 LeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGluLysIleValGlyArgTyrLeuGly 189  
Db 448 CTGGCTACGGCTTCTACGGGACATCATCAGGACAGTGAAGAACCGTGTGGT 507  
QY 190 LeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisGlyCysTyrGluGly 209  
Db 508 CTGGCAGATACGACTTTTCAGGTTTAAAGACTTCTCTCCACCACTGCTATGAAGGG 567  
QY 210 ThrValSerPheLeu 214  
Db 568 ACAGTGTCTCTCTC 582

BP309990 584 bp mRNA linear EST 17-SEP-2004  
BP309990 Sugano cDNA library, brain Homo sapiens cDNA clone  
NR01216, mRNA sequence.

ACCESSION BP309990

VERSION BP309990.1 GI:52238965

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Suzuki,Y., Yanashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,

Mizushima-Sugano,J., Nakai,K. and Sugano,S.

TITLE Sequence comparison of human and mouse genes reveals a homologous

block structure in the promoter regions

JOURNAL Genome Res. 14 (9), 1711-1718 (2004)

COMMENT Contact: Yuraka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: ysuzuki@ims.u-tokyo.ac.jp.

FEATURES Location/Qualifiers

1..584

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="NR01216"

/tissue\_type="brain"

/clone\_lib="Sugano cDNA library, brain"

ORIGIN

Alignment Scores: 3,42e-96 Length: 584  
Pred. No.: 965.50 Matches: 191  
Score: 93.69% Conservative: 2  
Percent Similarity: 92.72% Mismatches: 2  
Best Local Similarity: 56.23% Indels: 12  
Query Match: 5 Gaps: 1  
DB: 5  
US-10-631-958-2 (1-326) x BP309990 (1-584)  
QY 10 ProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPhe 29  
Db 1 CCGTTTGAGGNAAGGAAA-GGCNAGCGGTATATGAAGAAAAGTGGCACCCTGTTTC 59  
QY 30 ThrLeuAlaSerIleThrThrAspIleIleGlyAsnLysPheTyrValAsnTyrValGlu 49  
Db 60 ACCTTAGCCTCCATCACCCTGACATCATC----- 89  
QY 50 ValIleThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLys 69  
Db 90 ---GTTACTGAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAA 146  
QY 70 TyrAspGlyIleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeu 89  
Db 147 TACGACGGCATGCTGTCTCGCGGAGATGGTATGTTTCAGCGAGGTGCTGCACGGTCTG 206  
QY 90 IleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuVal 109  
Db 207 ATTGGAGAGCGCANAGGAGCGCGGGTCCACCAAGAACCCCGGGCTGTGTCTG 266  
QY 110 ProSerSerLeuArgIleGlyIleIleProAlaGlySerThrAspCysValCysTyrSer 129  
Db 267 CCCAGTAGCCTCCGATTCGATCATTCCTCCGAGGGTCAACGGACTGCGTGTACTCC 326  
QY 130 ThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeu 149  
Db 327 ACCGTGGGACCAAGCAGCAGAACCTTCGGGCTGCATATCGTTGTTGGGACTCGCTG 386  
QY 150 AlaMetAspValSerSerValHisAsnSerThrLeuLeuArgTyrSerValSerLeu 169  
Db 387 GCATGGAGTGTCTCAGTCACCAACAGCACTCCTTCGCTACTCCGTGTCTCTG 446  
QY 170 LeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGluLysIleValGlyArgTyrLeuGly 189  
Db 447 CTGGCTACGGCTTCTACGGGACATCATCAGGACAGTGAAGAACCGTGTGGT 506  
QY 190 LeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisGlyCysTyrGluGly 209  
Db 507 CTTCAGATACGACTTTTCAGGTTTAAAGACTTCTCTCCACCACTGCTATGAAGGG 566  
QY 210 ThrValSerPheLeuPro 215  
Db 567 ACAGTGTCTCTCTCCTCCCT 584  
RESULT 13  
CF138275  
LOCUS  
DEFINITION UI-HF-BNO-anz-g-11-0-UI.r1 NIH\_MGC\_50 Homo sapiens cDNA clone  
IMAGE:3094461 5', mRNA sequence.  
ACCESSION CF138275  
VERSION CF138275.1 GI:33253719  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 573)  
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 9704477



Db 184 TACTCGTGTCCGTGCTGGCTACGCGTTCTACGGGACATCATCAAGGACAGTGAAG 243

Qy 185 LysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHis 204

Db 244 AAACGGTGGTGGGTCTTCCAGATACGACTTTTCAGGTTTAAAGACCTTCTCTCCAC 303

Qy 205 HisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArg 224

Db 304 CACTGCTATGAAGGACAGACTGCTCTCCCTCGCACACACACGCGTGGATCTCCAAG 363

Qy 225 AspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnLeuGlu 244

Db 364 GATAGGAAGCCCTGCCGGGACAGATGCTTTGTTTCAGGCAAGCAAGCAGCAGTGGAG 423

Qy 245 GluGluGlnLysIleValLeuTyrGlyLeuGluAlaAlaGluAspValGluGlnTrpGln 264

Db 424 GAGGAGCAGAAAGACACTGTATGGTTTGAAGCTGCGAGGAGCTGGAGGAGTGGCAA 483

Qy 265 ValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArg 283

Db 484 GTTCTGTGGGAAGTTTCTGCCATCAATGCCAACCAAAATGCTCTGTGCTTGTGCGC 540

## RESULT 15

BU214294 603756937F1 CSEQCHN04 Gallus gallus cDNA clone ChEST669i10 5', mRNA

## LOCUS

603756937F1 CSEQCHN04 Gallus gallus cDNA clone ChEST669i10 5', mRNA

## DEFINITION

sequence.

## ACCESSION

BU214294

## VERSION

BU214294.1 GI:25390255

## KEYWORDS

EST.

## SOURCE

Gallus gallus (chicken)

## ORGANISM

Gallus gallus

## REFERENCE

1 (bases 1 to 725)

Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,

Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.

A Comprehensive Collection of Chicken cDNAs

Curr. Biol. 12 (22), 1965-1969 (2002)

22335534

12445392

Contact: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology

(UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

Location/Qualifiers

1. .725

/organism="Gallus gallus"

/mol\_type="mRNA"

/strain="White Leghorn, Hisex"

/db\_xref="taxon:9031"

/clone="ChEST669i10"

/tissue\_type="whole embryo"

/dev\_stage="20-21"

/lab\_host="DH10B"

/clone\_lib="CSEQCHN04"

/note="Organ: whole embryo; Vector: pBluescript II KS(+);

Site 1: EcoRI; Site 2: NotI; This normalized library was

constructed from 1 million independent clones. cDNA

synthesis was initiated using an oligo(dT) primer, using

methylated C in the first strand synthesis reaction.

Following this first strand reaction, double-stranded cDNA

was blunt-ended, ligated to NotI adapters, digested with

EcoRI, size-selected, and cloned into the NotI and EcoRI

compatible sites of a custom modified MCS of the

pBluescript (KS+) vector. The library was normalized in 2

rounds using conditions adapted from Soares et al., PNAS

(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6

(1996) 6: 791, except that a significantly longer

reannealing hybridization was used."

ORIGIN

Alignment Scores:

Pred. No.: 1.45e-91 Length: 725

Score: 925.00 Matches: 168

Percent Similarity: 84.84% Conservative: 39

Best Local Similarity: 68.85% Mismatches: 33

Query Match: 53.87% Indels: 4

DB: 5 Gaps: 1

US-10-631-958-2 (1-326) x BU214294 (1-725)

Qy 79 AspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGly 98

Db 3 GACGCGATGTTTCAGTGAAGTGTGCTCATTTGGAAGAATGCAAGAGCACTCTGGC 62

Qy 99 ValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIle 118

Db 63 ATAGACCAAAATAATCCCAAGACCCGTTAGTCCAGTCAATATAAGGATTTGGCAAT 122

Qy 119 ProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThr 138

Db 123 CCTGCTGCTCAACAGATGCTGATGCTATTCACACTGTTGGCATTTCTGATCCAGTAACA 182

Qy 139 SerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHis 158

Db 183 TCAGCTCTTCATATTTATATAGGTGACTGTGAGCCCTTTAGATGCTCATCTGTGCATCAA 242

Qy 159 AsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIle 178

Db 243 AACACACGATTTTGAAGTATGCTGTATCATTTGTTGGGCTATGTTTATGGAGATGTT 302

Qy 179 IleLysAspSerGluLysLysLysArgTrpLeuLysLysLysLysLysLysLysLys 198

Db 303 TTGAAGACATGAAGAAGAACCGTGGATGGTCCGATGAGATACGACTATTTCAGGCTTC 362

Qy 199 LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis 218

Db 363 AAGACTTTCTTCTCATCTACTATGAGGAAACAATTTCTTTTCAACACGACAAACAC 422

Qy 219 ThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGln 238

Db 423 ACCTGGGATCTCCACGAGATAAAGATAGCTGCAGAACAGGATGTTATTTTGAAGAA 482

Qy 239 SerLysGlnGlnLeuGluGluGlnLysLysLysLysLysLysLysLysLysLys 258

Db 483 AGTGGCGACAACTGGCGAGAACACGCAAGAAAGTGT-----GGATTCAAACATGAAGAA 536

Qy 259 AspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMet 278

Db 537 GATGAAGAAGAAATGGAAGGTTATTACCGGGAAATTTCTAGCCATCAATGTAGTAAATATG 596

Qy 279 SerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGly 298

Db 597 TGCTGTGCTGTCCACGAAGTCCAAAGGCTCTTTCCCGGAGCTCATTTTACGAGATGGT 656

Qy 299 SerSerAspLeuIleLeuIleArg-LysCysSerArgPheAsnPheLeuArg-PheLeu 318

Db 657 TCAGTGCACCTGATTTAGTTCGGGAAATGCTCCAGATTTGATTTTCTGCGTTTATCTTG 716

Qy 318 LeArgHis 320

Db 717 TCAGACAT 724

Search completed: September 6, 2005, 20:29:11

Job time : 3154.86 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 04:25:44 ; Search time 60.6246 Seconds  
(without alignments)  
2753.633 Million cell updates/sec

Title: US-10-631-958-2  
Perfect score: 1717  
Sequence: 1 PKHLVFNPGKGQGRRI.....KSRFNFLRFLRHTNQDQ 326

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1640.5	95.5	537	1 CEK1 HUMAN	Q8tc0 homo sapien
2	1503.5	87.6	409	2 Q62PK5	Q62pk5 mus musculus
3	1503.5	87.6	531	1 CEK1 MOUSE	Q8k4q7 mus musculus
4	1295	75.4	339	2 Q6NKS59	Q6nx59 homo sapien
5	1140	66.4	485	2 Q6GMF3	Q6gmf3 xenopus lae
6	1117.5	65.1	572	2 Q6GLV1	Q6glv1 xenopus lae
7	620.5	36.1	410	2 Q7PRA8	Q7pra8 anopheles g
8	613	35.7	487	2 Q95S15	Q95s15 drosophila
9	613	35.7	687	2 Q9VNA6	Q9vna6 drosophila
10	502.5	29.3	533	2 Q9LU45	Q9lu45 arabidopsis
11	498.5	29.0	608	2 Q6USK2	Q6usk2 arabidopsis
12	486.5	28.3	549	2 Q9T211	Q9t2i1 caenorhabdi
13	455.5	26.5	700	2 Q6H6H1	Q6h6h1 oryza sativ
14	436	25.4	532	2 Q6UZF6	Q6ufz6 homo sapien
15	389	22.7	485	2 Q817L1	Q817l1 arabidopsis
16	365.5	21.3	1240	2 Q65419	Q65419 arabidopsis
17	321.5	18.7	586	2 Q949C3	Q949c3 oryza sativ
18	311.5	18.1	732	2 Q9PHL3	Q9phl3 arabidopsis
19	311.5	18.1	763	2 Q9LRB0	Q9lrbo arabidopsis
20	307.5	17.9	480	2 Q8H350	Q8h350 oryza sativ
21	301	17.5	685	2 Q7Q1P4	Q7q1p4 anopheles g
22	300	17.5	446	2 Q84S01	Q84s01 oryza sativ
23	292	17.0	748	2 Q7XN57	Q7xn57 oryza sativ
24	287.5	16.7	757	2 Q94HY9	Q94hy9 oryza sativ
25	287.5	16.7	757	2 Q7KCS9	Q7kcs9 oryza sativ
26	282	16.4	641	2 Q9VYX8	Q9vyx8 drosophila
27	268	15.6	661	2 Q9VZW0	Q9vzw0 drosophila
28	261	15.2	280	2 Q6ZP59	Q6zp59 homo sapien
29	244.5	14.2	624	2 Q86KF9	Q86kf9 dictyosteli
30	243	14.2	654	1 SPH2 MOUSE	Q8nra0 homo sapien
31	236	13.7	616	2 Q6AYB2	Q6ayb2 rattus norv

32	236	13.7	517	1 SPH2 MOUSE	Q9jia7 mus musculus
33	231	13.5	579	2 Q74E3	Q74ez3 ashbya goss
34	227.5	13.2	384	1 SPH1 HUMAN	Q9nyal homo sapien
35	227.5	13.2	384	2 Q96HV8	Q96hv8 homo sapien
36	227.5	13.2	398	2 Q96GK1	Q96gk1 homo sapien
37	227.5	13.2	470	2 Q8N632	Q8n632 homo sapien
38	226	13.2	381	2 Q91ZN3	Q91zn3 mus musculus
39	226	13.2	382	2 Q8CI15	Q8cil5 mus musculus
40	226	13.2	388	2 Q88886	Q88886 mus musculus
41	226	13.2	458	2 Q14159	Q14159 schizosacch
42	226	13.2	504	2 Q88885	Q88885 mus musculus
43	225.5	13.1	504	2 Q6BW14	Q6bw14 debaryomyce
44	224	13.0	329	2 Q9N0A5	Q9n0a5 macaca fasc
45	220.5	12.8	687	2 Q06147	Q06147 saccharomyc

#### ALIGNMENTS

RESULT 1  
CEK1\_HUMAN STANDARD; PRT; 537 AA.  
ID AC Q8TCT0; Q9BYB3; Q9UGE5;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Ceramide kinase (EC 2.7.1.138) (Acylsphingosine kinase) (hCERK) (Lipid kinase 4) (LK4).  
GN Name=CERK; Synonyms=KIAA1646;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
RC TISSUE=Leukemia;  
RX MEDLINE=22075121; PubMed=11956206; DOI=10.1074/jbc.M201535200;  
RA Sugiyama M., Kono K., Liu H., Shimizugawa T., Minekura H., Spiegel S., Kohama T.;  
RT "Ceramide kinase, a novel lipid kinase. Molecular cloning and functional characterization."  
RL J. Biol. Chem. 277:23294-23300(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Van Veldhoven P.P.;  
RT "A search for lipid kinases."  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20057165; PubMed=10591208; DOI=10.1038/990031;  
RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M., Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K., Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J., Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G., Clegg S.W., Cobley V.E., Cole C.G., Collier R.E., Connor R., Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E., Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G., Evans K.L., Fey J.M., Fleming K., French L., Garner A.A., Gilbert J.G.R., Goward M.E., Graham D.V., Griffiths M.N.D., Hall C., Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S., Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A., Laird G.K., Langford C.F., Levergha M.A., Lloyd C., Lloyd D.M., Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., McCann O.T., McClay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J., Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T., Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Rose M.T., Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L., Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M., Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L., Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L., Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N., Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,

RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,  
 RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,  
 RA Dorwan A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,  
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H., I.,  
 RA Wang Q., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,  
 RA Phang S., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,  
 RA Zhan M., Zhang G., Chisoso S., Murray J., Miller N., Minx P.,  
 RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,  
 RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,  
 RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,  
 RA Schreep P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,  
 RA Korf I., Bedell J.A., Hillier L.W., Marsis E., Waterston R.,  
 RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,  
 RA Edelard M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,  
 RA Edelman L., Kim U.J., Shiruya H., Simon M.I., Dumaneki J.P.,  
 RA Peyrard M., Kedra D., Seroussi E., Fransson I., Tapia I., Bruder C.E.,  
 RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,  
 RA Khan A.S., Lane L., Tilahun Y., Wright H.,  
 RA "The DNA sequence of human chromosome 22."  
 RL Nature 402:489-495(1999).  
 [4]  
 RN SEQUENCE OF 57-537 FROM N.A.  
 RP TISSUE=Brain;  
 RC MEDLINE=21156230; PubMed=11258795;  
 RX Hirasawa M., Nagase T., Murahashi Y., Kikuno R., Ohara O.;  
 RA "Identification of novel transcribed sequences on human chromosome 22  
 RT by expressed sequence tag mapping."  
 RL DNA Res. 8:1-9(2001).  
 CC -!- FUNCTION: Catalyzes specifically the phosphorylation of ceramide  
 CC to form ceramide 1-phosphate. Acts efficiently on natural and  
 CC analog ceramides (C6, C8, C16 ceramides, and C8-dihydroceramide),  
 CC to a lesser extent on C2-ceramide and C6-dihydroceramide, but not  
 CC on other lipids, such as various sphingosines.  
 CC -!- CATALYTIC ACTIVITY: ATP + ceramide = ADP + ceramide 1-phosphate.  
 CC -!- COFACTOR: Calcium and magnesium.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.  
 CC -!- TISSUE SPECIFICITY: High level expression in heart, brain,  
 CC skeletal muscle, kidney and liver; moderate in peripheral blood  
 CC leukocytes and thymus; very low in spleen, small intestine,  
 CC placenta and lung.  
 CC -!- MISCELLANEOUS: Optimal pH is 6.0-7.5.  
 CC -!- SIMILARITY: Contains 1 DAGKC domain.  
 CC -!- CAUTION: Ref.3 sequence differs from that shown due to erroneous  
 CC gene model prediction. An additional exon may exist between amino  
 CC acid positions 168 and 169.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AB079066; BAC01154.1; -;  
 DR EMBL; AJ457828; CAD29884.1; -;  
 DR EMBL; AL096766; CAB62977.1; ALT SEQ.  
 DR EMBL; AL118536; -; NOT ANNOTATED\_CDS.  
 DR EMBL; AB051433; BAB3316.1; -;  
 DR EMBL; HGNC:19256; CERK.  
 DR GO; GO:0000299; C:integral to membrane of membrane fraction; IDA.  
 DR GO; GO:0004685; F:calcium/calmodulin-dependent protein kinase. . .; IDA.  
 DR GO; GO:0001729; F:ceramide kinase activity; IDA.  
 DR GO; GO:0000287; F:magnesium ion binding; IDA.  
 DR GO; GO:0006672; F:ceramide metabolism; TAS.  
 DR InterPro; IPR001206; DAGKC.  
 DR InterPro; IPR001849; PH.  
 DR InterPro; IPR011036; PH related.  
 DR Pfam; PF00781; DAGK\_cat; 1.  
 DR ProDom; PD005043; DAGKC; 1.  
 DR SMART; SM00046; DAGKC; 1.  
 DR SMART; SM00233; PH; 1.  
 KW Calcium; Kinase; Magnesium; Transferase.

FT DOMAIN 132 278 DAGKC.  
 SQ SEQUENCE 537 AA; 59977 MW; 3DBFC0ED8D679F7F CRC64;  
 Query Match 95.5%; Score 1640.5; DB 1; Length 537;  
 Best Local Similarity 96.3%; Pred. No. 8.7e-135; Indels 11; Gaps 1;  
 Matches 314; Conservative 1; Mismatches 0;  
 QY 1 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTTDIIGNKFFVNVYVEITEHANQAKE 60  
 Db 130 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTTDI-----VTEHANQAKE 178  
 QY 61 TLVEINIDKDYGVVCGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSLSLRIGIIPA 120  
 Db 179 TLVEINIDKDYGVVCGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSLSLRIGIIPA 238  
 QY 121 GSTDCVCYSTVGTSDAETSAHIVVGDLSAMDVSSVHHNSTLLRYSYSLSLGCGFVGDIK 180  
 Db 239 GSTDCVCYSTVGTSDAETSAHIVVGDLSAMDVSSVHHNSTLLRYSYSLSLGCGFVGDIK 298  
 QY 181 DSEKKRWGLARYDFSGIKTFLSHHCYEGTVSFPLPAQHTVGSPPDRKPCRAGCFVCQSK 240  
 Db 299 DSEKKRWGLARYDFSGIKTFLSHHCYEGTVSFPLPAQHTVGSPPDRKPCRAGCFVCQSK 358  
 QY 241 QOLEEEOKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300  
 Db 359 QOLEEEOKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 418  
 QY 301 DLILIRKCSRFNLFIRHTNQDDQ 326  
 Db 419 DLILIRKCSRFNLFIRHTNQDDQ 444  
 RESULT 2  
 Q6ZPK5 PRELIMINARY; PRT; 409 AA.  
 ID Q6ZPK5 AC Q6ZPK5; DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE MKTAA1646 protein (Fragment).  
 GN Name=mKIAA1646;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_TaxID=10090;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryonic tail;  
 RX PubMed=14621295;  
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,  
 RA Saga Y., Nagase T., Ohara O., Koga H.;  
 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:  
 RT cDNAs identified by screening of terminal sequences of 500 mouse KIAA-homologous  
 RT randomly sampled from size-fractionated libraries."  
 RL DNA Res. 10:167-180(2003).  
 DR EMBL; AK129416; BAC98226.1; -;  
 DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.  
 DR GO; GO:0007205; P:protein kinase C activation; IEA.  
 DR InterPro; IPR001206; DAGKC.  
 DR Pfam; PF00781; DAGK\_cat; 1.  
 DR ProDom; PD005043; DAGKC; 1.  
 DR SMART; SM00046; DAGKC; 1.  
 FT NON TER 1  
 SQ SEQUENCE 409 AA; 45697 MW; 810E815B37E26808 CRC64;  
 Query Match 87.6%; Score 1503.5; DB 2; Length 409;  
 Best Local Similarity 86.8%; Pred. No. 6e-123;  
 Matches 283; Conservative 19; Mismatches 13; Indels 11; Gaps 1;  
 QY 1 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTTDIIGNKFFVNVYVEITEHANQAKE 60  
 Db 8 PKHLLVFNPFGGKGQKRIYERKVAFLFTLASITTTII-----ITEHANQAKE 56

QY 61 TLYEINIDKYDGVVCGDGMFSEVLHGLIGTQRSAGVDQNHPRVLPVPSLRIGIIPA 120  
 DB 57 TLYEINTDSYDGVVCGDGMFSEVLHGLIGTQRSAGVDQNHPRVLPVPSLRIGIIPA 116  
 QY 121 GSTDCVCVSTGCTSDAETSALHIVGDSLMDVSSVHNSLTLRYSVLLGFGFYGDILK 180  
 DB 117 GSTDCVCVSTGCTSDAETSALHIVGDSLMDVSSVHNSLTLRYSVLLGFGFYGDILK 176  
 QY 181 DSEKRWGLGARYDFSGLTFLSHHCYEGTGVFLPAQHTVGSPPDRKPCRCAGFCVCRQSK 240  
 DB 177 DSEKRWGLGARYDFSGLTFLSHHCYEGTGVFLPAQHTVGSPPDRKPCRCAGFCVCRQSK 236  
 QY 241 QOLEEEOKKALYGLAEAEDEVEWQVCGKFLAINATNNSCACRSPRGLSPAHLGDGSS 300  
 DB 237 QOLEEEOKKALYGLAEAEDEVEWQVCGKFLAINATNNSCACRSPRGLSPAHLGDGSS 296  
 QY 301 DLILIRKSRFNFRLIRHTNQDQ 326  
 DB 297 DLILIRKSRFNFRLIRHTNQDQ 322

RESULT 3  
 CEK1\_MOUSE  
 ID CEK1\_MOUSE STANDARD; PRT; 531 AA.  
 AC Q8K407;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Ceramide kinase (EC 2.7.1.138) (Acylephingosine kinase) (mCKERK).  
 GN Name=Cerk;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=22075121; PubMed=11956206; DOI=10.1074/jbc.M201535200;  
 RA Sugitara M., Kono K., Liu H., Shimizugawa T., Minekura H., Spiegel S.,  
 RA Kohama T.;  
 RA "Ceramide kinase, a novel lipid kinase. Molecular cloning and  
 RT functional characterization.";  
 RL J. Biol. Chem. 277:23294-23300 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Heart, and Thymus;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Balderelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuoka H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Perte G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reid D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
 RA Sultana R., Tatenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shikata T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Iehii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki K., Shibata K., Shingawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,

RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573 (2002).  
 CC -!- FUNCTION: Catalyzes specifically the phosphorylation of ceramide  
 CC to form ceramide 1-phosphate. Acts efficiently on natural and  
 CC analog ceramides (C6, C8, C16 ceramides, and C8-dihydroceramide),  
 CC to a lesser extent on C2-ceramide and C6-dihydroceramide, but not  
 CC on other lipids, such as various sphingosines (By similarity).  
 CC -!- CATALYTIC ACTIVITY: ATP + ceramide = ADP + ceramide 1-phosphate.  
 CC -!- COFACTOR: Calcium and magnesium (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By  
 CC similarity).  
 CC -!- TISSUE SPECIFICITY: High level expression in heart, brain and  
 CC testis; low expression in spleen, liver and lung; not detected in  
 CC skeletal muscle.  
 CC -!- DEVELOPMENTAL STAGE: Highly expressed at 7 dpc and decreases  
 CC rapidly thereafter.  
 CC -!- SIMILARITY: Contains 1 DAGKc domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AB079067; BAC01155.1; -;  
 DR EMBL; AK042077; BAC31157.1; -;  
 DR EMBL; AK052269; BAC34908.1; -;  
 DR MGI; MGI:2386052; Cerk.  
 DR GO; GO:0000299; C:integral to membrane of membrane fraction; ISS.  
 DR GO; GO:0004685; F:calcium/calmodulin-dependent protein kinase. . ; ISS.  
 DR GO; GO:0001729; F:ceramide kinase activity; ISS.  
 DR GO; GO:0000287; F:magnesium ion binding; ISS.  
 DR GO; GO:0006672; P:ceramide metabolism; ISS.  
 DR InterPro; IPR001206; DAGKc.  
 DR InterPro; IPR011036; PH-related.  
 DR Pfam; PF00781; DAGK\_cat; 1.  
 DR ProDom; PD005043; DAGKc; 1.  
 DR Domain; Kinase; Magnesium; Transferase.  
 FT DOMAIN 132 278  
 FT CONFLICT 378 378 V -> M (in Ref. 2).  
 FT CONFLICT 467 467 D -> Y (in Ref. 2; BAC34908).  
 SQ SEQUENCE 531 AA; 59811 MW; 31FEC534C348AA0A CRC64;  
 Query Match 87.6%; Score 1503.5; DB 1; Length 531;  
 Best Local Similarity 86.8%; Pred. No. 8.4e-123;  
 Matches 283; Conservative 19; Mismatches 13; Indels 11; Gaps 1;  
 QY 1 PKHLVFNPFPGKGQGRIVYERKVAFLFTLASITTDIIGNKFYVYVEVITEHANOAKE 60  
 DB 130 PKHLVFNPFPGKGQGRIVYERKVAFLFTLASITTEII-----ITEHANOAKE 178  
 QY 61 TLYEINIDKYDGVVCGDGMFSEVLHGLIGTQRSAGVDQNHPRVLPVPSLRIGIIPA 120  
 DB 179 TLYEINTDSYDGVVCGDGMFSEVLHGLIGTQRSAGVDQNHPRVLPVPSLRIGIIPA 238  
 QY 121 GSTDCVCVSTGCTSDAETSALHIVGDSLMDVSSVHNSLTLRYSVLLGFGFYGDILK 180  
 DB 239 GSTDCVCVSTGCTSDAETSALHIVGDSLMDVSSVHNSLTLRYSVLLGFGFYGDILK 298  
 QY 181 DSEKRWGLGARYDFSGLTFLSHHCYEGTGVFLPAQHTVGSPPDRKPCRCAGFCVCRQSK 240  
 DB 299 DSEKRWGLGARYDFSGLTFLSHHCYEGTGVFLPAQHTVGSPPDRKPCRCAGFCVCRQSK 358  
 QY 241 QOLEEEOKKALYGLAEAEDEVEWQVCGKFLAINATNNSCACRSPRGLSPAHLGDGSS 300  
 DB 359 QOLEEEOKKALYGLAEAEDEVEWQVCGKFLAINATNNSCACRSPRGLSPAHLGDGSS 418  
 QY 301 DLILIRKSRFNFRLIRHTNQDQ 326  
 DB 297 DLILIRKSRFNFRLIRHTNQDQ 322



```
Db 419 DLILIRKCSRFNLFRLIRHTNQEDQ 444
RESULT 4
Q6NX59
ID Q6NX59 PRELIMINARY; PRT; 339 AA.
AC Q6NX59
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE CERK protein.
GN Name=CERK;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral Nervous System;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Touchman A.C., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral Nervous System;
RX Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL: BC067255; AAH67255.1; -.
DR GO: GO:0004143; F:diacylglycerol kinase activity; IEA.
DR GO: GO:0007205; P:protein kinase C activation; IEA.
DR InterPro: IPR001206; DAGKC.
DR ProDom: PD005043; DAGKC; 1.
SQ SEQUENCE 339 AA; 37780 MW; A4C2ACDFF2E6F3D0 CRC64;

Query Match 75.4%; Score 1295; DB 2; Length 339;
Best Local Similarity 100.0%; Pred. No. 8.4e-105;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 MFSEVLHGLIGRTORSAGVDQNHPRVLPVPSLSRLGIIIPAGTDCVCYSTVTGTSDAETSA 140
Db 1 MFSEVLHGLIGRTORSAGVDQNHPRVLPVPSLSRLGIIIPAGTDCVCYSTVTGTSDAETSA 60

Qy 141 LHIVVGDSLAMDVSSVHNSLIRYSVSLGGLGYFGYDIIKDSEKKRWGLGLARYDFSGLKT 200
Db 61 LHIVVGDSLAMDVSSVHNSLIRYSVSLGGLGYFGYDIIKDSEKKRWGLGLARYDFSGLKT 120

Qy 201 FLSHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCROSKQOLEBEQKALYGLEAARDV 260
Db 121 FLSHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCROSKQOLEBEQKALYGLEAARDV 180

Qy 261 EEWQVCCGFLAINATNNSCARRSPRGISPAAHLDGSSDLILIRKCSRFNLFRLIRH 320
Db 181 EEWQVCCGFLAINATNNSCARRSPRGISPAAHLDGSSDLILIRKCSRFNLFRLIRH 240
-Qy 321 TNQDDQ 326
```

```
Db 241 TNQDDQ 246
|||||
|||||
RESULT 5
Q6GMF3
ID Q6GMF3 PRELIMINARY; PRT; 485 AA.
AC Q6GMF3
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE MGC81777 protein.
GN Name=MGC81777;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Touchman A.C., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative."
RL Dev. Dyn. 225:384-391 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Gerhardt D.S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC074110; AAH74110.1; -.
DR GO: GO:0004143; P:diacylglycerol kinase activity; IEA.
DR GO: GO:0007205; P:protein kinase C activation; IEA.
DR InterPro: IPR001206; DAGKC.
DR ProDom: PD005043; DAGKC; 1.
DR SMART: SM00046; DAGKC; 1.
SQ SEQUENCE 485 AA; 54874 MW; E7C30C87AE52214F CRC64;

Query Match 66.4%; Score 1140; DB 2; Length 485;
Best Local Similarity 62.4%; Pred. No. 4.9e-91;
Matches 204; Conservative 58; Mismatches 53; Indels 12; Gaps 2;

Qy 1 PKHLVFINPGGKGGCKRIYERKVLFTLASITTDIIGNKFFVNVYVEVITEHANOAKE 60
Db 157 PKYLLVINPYGKGGKQIYETKVLFLSAAAGICADV-----VTEYANHARD 205
```

```
QY 61 TLYEINIDKYDGIIVCGDGMSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPA 120
DB 206 NUYEWNLEKYDGIIVCGDGMSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPA 265
QY 121 GSTDCVCYSTGTSDAETSSALHIVVGDLSLMDVSVVHNSITLLRYSVSLGFGFYGDIIK 180
DB 266 GSTDCICVATGINDPETSALHIIIGDCQPLDVCVSHVYKRTFLKYVSLGFGFYGDVLK 325
QY 181 DSEKRWGLGARYDPSGLKTFLSHHCHYGTGVSFLPAQHTVGSPPDRKPCRCAGCFVCROSK 240
DB 326 GTEKRWGLGARYDPSGCKTFLTHHCYGSVSFPQAKWLVGSPPDRQTPCTSGCYICROSS 385
QY 241 QOLEEEOKKALYGLE-AEDVEEVMQVCGKFLAINATNMSCACRSPRGLSPAHLGDS 299
DB 386 KOLEEQKQTVGSGEHRQDDWTITIGRFMAINAVSMSCACRPTPKGLSPAHLADGS 445
QY 300 SDLIIRKCSRFNFLRLIRHTNQDQ 326
DB 446 ADLIIVKCSRLDFLRLIRHTSNKDQ 472

RESULT 6
Q6GLV1
ID Q6GLV1 PRELIMINARY; PRT; 572 AA.
AC Q6GLV1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE MGC84197 protein.
GN Name=MGC84197;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2];
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.,
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3];
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Klein S., Gerhard D.S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; BC074350; AAH74350.1; -
DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
DR InterPro; IPR001206; DAGKC.
DR Pfam; PF00781; DAGK_cat; 1.
DR ProDom; PD005043; DAGKc; 1.
DR SMART; SM00046; DAGKc; 1.
SQ SEQUENCE 572 AA; 64899 MW; E3D22DB2C93B5B77 CRC64;

Query Match 65.1%; Score 1117.5; DB 2; Length 572;
Best Local Similarity 62.0%; Pred. No. 5.6e-89;
Matches 202; Conservative 54; Mismatches 59; Indels 11; Gaps 1;

QY 1 PKHLVFINPFGGKGQKRIYERKVAFLFTLASITTDIIGNKFFYNYVEVTEHANOAKE 60
DB 157 PKNLVFINPFGGKGQKRIYERKVAFLFTLASITTDIIGNKFFYNYVEVTEHANOAKE 205
QY 61 TLYEINIDKYDGIIVCGDGMSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPA 120
DB 206 HLYDANLEKYDGIIVCGDGMSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPA 265
QY 121 GSTDCVCYSTGTSDAETSSALHIVVGDLSLMDVSVVHNSITLLRYSVSLGFGFYGDIIK 180
DB 266 GSTDCICVATGINDPETSALHIIIGDCQPLDVCVSHVYKRTFLKYVSLGFGFYGDVLK 325
QY 181 DSEKRWGLGARYDPSGLKTFLSHHCHYGTGVSFLPAQHTVGSPPDRKPCRCAGCFVCROSK 240
DB 326 GSEKRWGLGARYDPSGCKTFLTHHCYGSVSFPQAKWLVGSPPDRQTPCTSGCYICROSS 385
QY 241 QOLEEEOKKALYGLEAEDVEEVMQVCGKFLAINATNMSCACRSPRGLSPAHLGDS 300
DB 386 KOLEEQKQTVGSGEHRQDDWTITIGRFMAINAVSMSCACRSPNGLSPAHLADGSA 445
QY 301 DLIIRKCSRFNFLRLIRHTNQDQ 326
DB 446 DLIIVKCSRLDFLRLIRHTSNKDQ 471

RESULT 7
Q7PRAB
ID Q7PRAB PRELIMINARY; PRT; 410 AA.
AC Q7PRAB;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENGANGP0000010845 (Fragment).
GN Name=ENGANGP0000010845;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AA801008059; EAA07556.2; -
DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
DR InterPro; IPR001206; DAGKC.
DR Pfam; PF00781; DAGK_cat; 1.
DR ProDom; PD005043; DAGKc; 1.
FT NON TER 1
FT NON TER 410
SQ SEQUENCE 410 AA; 45669 MW; 79535E4C0D150154 CRC64;

Query Match 36.1%; Score 620.5; DB 2; Length 410;
Best Local Similarity 39.2%; Pred. No. 1.1e-45;
Matches 135; Conservative 53; Mismatches 119; Indels 37; Gaps 5;
```

```
Qy 1 PKHLLVFNPFGGKGQKRIYERKVPALFTLASITTDIIGNKFYVYVVEVITEHANOAKE 60
Db 77 PKHLLLFNPGGKNALARYERAKPLFRLAGVDINLI-----ITQRAQQIYD 125
Qy 61 --TLYEINIDKYDGVCGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGII 118
Db 126 IVTSKSLIDNDYGLVCCGGDTFAELFNGLVTRTMDMCGIDIKYP-AYLPKPNIPIGVI 184
Qy 119 PAGSTDVCYVSTGSDAETSAHLHVCGDSLAMDVSSVHH-----N 159
Db 185 PAGSTDVACCLNGTITDKTICIIILIQHSGLDISAVYSDAAKCDEGASPACTGRPR 244
Qy 160 STLLRYSVSLGYPGDIIDKSEKRWLGLARYDFGLKTLFSLHCHYEGTVSFLPAQHT 219
Db 245 PQLKLFPASALSYGLGDIAYDSEKYRWGPKRYDYGFKFLANRGYNABIVVHLDRRG 304
Qy 220 VGSFDRKPCRAGCFVCRSQKQLEBEOKKALYLEAAEDVEWQVCGKFLAINATNMS 279
Db 305 KQDPNDGVRCLDKCARCKKAKYGRDCGGERASY-----EDDTEPLVVRGKFLMVSGANIS 360
Qy 280 CACRRSPRGLSPAHLGDSGLILIRKCSRFNLFRLIRHTNQ 323
Db 361 CSCERSPOGFSYCHLGDGLDLVLVRHTSFNMLRLLLTWTMSK 404

RESULT 8
Q9VNA6
ID Q9VNA6 PRELIMINARY; PRT; 487 AA.
AC Q9VNA6;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE HL015389.
GN ORFNames=CG16708;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunco J., Paclab J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY061001; AAL28549.1; -.
DR FlyBase; FBgn0037315; CG16708.
DR GO; GO:0048102; P:autophagic cell death; IEP.
DR InterPro; IPR001206; DAGKC.
DR Pfam; PF00781; DAGK_cat; 1.
DR SMART; SM00046; DAGKC; 1.
SQ SEQUENCE 487 AA; 53711 MW; A4E71EC40354BB07 CRC64;

Query Match 35.7%; Score 613; DB 2; Length 487;
Best Local Similarity 35.2%; Pred. No. 6.5e-45;
Matches 134; Conservative 59; Mismatches 108; Indels 80; Gaps 7;

Qy 2 KHLVFNPFPGKGQKRIYERKVPALFTLASITTDIIGNKFYVYVVEVITEHANOAKE 61
Db 4 RLLVFNFPYGGKAGATYERHVRPIFQLAGVDATCI-----TTQRANQVKDI 52
Qy 62 LYEINIDKYDGVCGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGIIPAG 121
Db 53 LLSHDLGVYDACCVCVGGDTVAEVLINGLI FRQRELGLDEQRPPIPRP-ALPVGVIIPAG 111
Qy 122 STDVCYVSTGSDAETSAHLHVCGDSLAMDVSSVHHNSTLLRYSVSLGFGFYDIIKD 181
Db 112 STDTIAYSMHGTAADVRTAAIHVLQHRGLDVCVSGSVSQSLRLRPFCAVSLSYGLGDVAQ 171
```

## RESULT 9

```
Q9VNA6
ID Q9VNA6 PRELIMINARY; PRT; 687 AA.
AC Q9VNA6;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE CG16708-PA (CG16708-pb).
GN ORFNames=CG16708;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=2019606; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Kocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Murtugulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert B.C., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RN Science 287:2185-2195(2000).
```



Q6USK2 Q6USK2 PRELIMINARY; PRT; 608 AA.  
AC Q6USK2;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Ceramide kinase.  
GN Name=CERK;  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22959668; PubMed=14563678; DOI=10.1101/gad.1140503;  
RA Liang H., Yao N., Song J.T., Luo S., Lu H., Greenberg J.T.;  
RT "Ceramide modulate programmed cell death in plants.";  
RL Genes Dev. 17:2636-2641(2003).  
DR EMBL; AY362552; AAQ62904.1; -.  
DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.  
DR GO; GO:0007205; P:protein kinase C activation; IEA.  
DR InterPro; IPR001206; DAGKC.  
DR Pfam; PF00781; DAGK\_cat; 1.  
DR ProDom; PD005043; DAGK; 1.  
KW Kinase.  
SQ SEQUENCE 608 AA; 68462 MW; 65EB9353692D9CA6 CRC64;  
  
Query Match 29.0%; Score 498.5; DB 2; Length 608;  
Best Local Similarity 33.6%; Pred. No. 9e-35;  
Matches 134; Conservative 48; Mismatches 112; Indels 105; Gaps 15;  
  
Qy 1 PKHLVFINPFGGKGQKRIYERKVAFLFTLASITTDIIGNKFYVNVVEVITEHANOAKE 60  
Db 162 PRNLVFPVHPSGKNGSKWE-TVSKIFIPRAKNTKVI-----VTERAGHAFD 209  
  
Qy 61 TLYEII---NIDKYDGIYVCGDGMFSEVLHG-LIGRTO-----RSAG---- 98  
Db 210 VMASIQNKELHTYDGIIVAGVGGDFNEITLNGYLLSLKVLPLPPSPDSFNSVQSRGSSV 269  
  
Qy 99 -----VDQ---NHPRAVLPSSL-----LIPDPSVQVEMNFRVNGSCEGIEDPHPFSSERPRFGLIP 326  
Db 270 PEPGDEVHETDQEHYP---LLPDSVQVEMNFRVNGSCEGIEDPHPFSSERPRFGLIP 326  
  
Qy 120 AGSTDVCYSTVGTSDAETSALHIVVGDLSAMDVSSV-----HNSLTU---LRYSVSLIG 171  
Db 327 AGSDAIVMCTTGARDPVTSAHLILGRKLFELDAMQVVRWKTASTSTIEPIRYAASFAG 386  
  
Qy 172 YGFYGDIIKSEKRWGLARYDFSGLKTFLSHHCYEGTVSFLPAQ-----HT---- 219  
Db 387 YGFYGDVISESEKYRWGPKRYDVGTGKIFLKHRSYEAEMFEEASENSKASLHTRSKT 446  
  
Qy 220 ---VGSPPDRK-PCRAGCFVCRSQKQLEEBQKALYLEAAEDVEEWQVVGKFLAINA 275  
Db 447 WPFRTTTRSEKILCRANKICNS-----KVGWNSASTTLNCPKTRWCRTKGRFLSIGA 501  
  
Qy 276 TNMSCACRRSPRGLSPAHLGGSDLLILIRKCSRFNPL 314  
Db 502 AVMSNRNERAPDGLVVDAAHSDGFLHLILIKDCSRPKYL 540  
  
RESULT 12  
Q9TZ11 Q9TZ11 PRELIMINARY; PRT; 549 AA.  
AC Q9TZ11;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein T10B11.2.  
GN Name=T10B11.2; ORFNames=T10B11.2;  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=Bristol N2;  
RC MEDLINE=99069613; PubMed=9851916;  
RG WormBase Consortium;  
RT "Genome sequence of the nematode C. elegans: a platform for  
RT investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=Bristol N2;  
RA Minx P., Kemp K.;  
RT "The sequence of C. elegans cosmid T10B11.";  
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX STRAIN=Bristol N2;  
RC Waterston R.;  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RX STRAIN=Bristol N2;  
RC Wilson R.;  
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RX STRAIN=Bristol N2;  
RG WormBase Consortium;  
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF098993; AAC67466.1; -.  
DR PIR; T33517; T33517.  
DR WormBase; WBGene00020398; T10B11.2.  
DR WormPep; T10B11.2; CE18241.  
DR InterPro; IPR001206; DAGKC.  
DR Pfam; PF00781; DAGK\_cat; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 549 AA; 62425 MW; DE9573755534EEB CRC64;  
  
Query Match 28.3%; Score 486.5; DB 2; Length 549;  
Best Local Similarity 34.6%; Pred. No. 8.9e-34;  
Matches 111; Conservative 66; Mismatches 113; Indels 31; Gaps 6;  
  
Qy 1 PKHLVFINPFGGKGQKRIYERKVAFLFTLASITTDIIGNKFYVNVVEVITEHANOAKE 60  
Db 164 PKNIIIFINPFGGNGKAQIKFNVDFAFWLTP-----GLRYKVVLTERRANHARD 213  
  
Qy 61 TLVEINTDKY---DGIYVCGDGMFSEVLHG-LIGRTOVSAGVDONHPRAVLPSSLRIGI 117  
Db 214 YIVEMPPQWSAIDGLVSGDGLFNELLSGALLRTQTDAGRINDPSSHLVTPHIRFGI 273  
  
Qy 118 IPAGSTDVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLIGYGFVGD 177  
Db 274 IGASANSIVSTVHETNDHATSVAHIAIGSECNVDVCTVHQHQLILISANAISYGLGD 333  
  
Qy 178 IIKDSEKRWGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPPDRK-KPCRAGCFVC 236  
Db 334 VLRSSEYRCIGPIRYQWSALRTIRHPYIRGMVQFSLSHKENVNPXQDLPPCLPCPVC 393  
  
Qy 237 RQSKQLEEBQKALYLEAAEDVEEWQVVGKFLAINATNMSCACRRSPRGLSPAHLG 296  
Db 394 MK-----PQGNDKYDYHWA-----EFTHVICCVIPTVTPF-----TPYGLAPFTGIG 436  
  
297 DGSDDLILIRKCSRFNPL 317  
437 DGTLDLALVPRISRFRHNNQFM 457  
  
RESULT 13  
Q6H6H1 Q6H6H1 PRELIMINARY; PRT; 700 AA.  
AC Q6H6H1;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DE Putative ceramide kinase.  
 OS Name=P0519806.23; Synonyms=OJ1003\_B06.39;  
 GN Oryza sativa (japonica cultivar-group)  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, PAC  
 clone:PD519E06";  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP005006; BAD25678.1; --  
 DR EMBL; AP004676; BAD25337.1; --  
 DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.  
 DR GO; GO:0007205; P:protein kinase C activation; IEA.  
 DR InterPro; IPR001206; DAGKC.  
 DR Pfam; PF00781; DAGK\_cat; 1.  
 DR ProDom; PD005043; DAGKC; 1.  
 KW Kinase.  
 SQ SEQUENCE 700 AA; 77961 MW; FBB2C2A3DF31C75D CRC64;  
 Query Match 26.5%; Score 455.5; DB 2; Length 700;  
 Best Local Similarity 29.8%; Pred. No. 6.3e-31;  
 Matches 134; Conservative 55; Mismatches 114; Indels 147; Gaps 16;  
 QY 1 PKHLVFINPFGKGQKRIYERKVAFLTLASITTDIIGNKFFYNYVEVITEHANQAKE 60  
 DB 182 PKSLAVFVHPLCGKRGCKNWE-TVAPLPERAKVTKVI-----VTQRAHAYD 229  
 QY 61 TLVEI---NIDKYDGIYCV-----  
 DB 230 TLASLDKDLKFPDGVIAVNTINACLSLFDIKHNYKMSARPENTLSYDPOSAASGHKSM 289  
 QY 77 -----GGDGMFSEVL-----HGLIGRTORSAGVD--- 100  
 DB 290 LIFYCFIINMKCOEHRNDLSNELTGDDANAISGSNTPDHPEFLSTTRSTGLDISS 349  
 QY 101 -----QNHPRALV--PSS-LRIGIIPAGSTDC-----VCYSTVGTSDAETSALHI 143  
 DB 350 SDSDDEPCNGDQPLVSPFNWFRLLIIPSGSTDAIVLSPVDVC-STTGERDPVTSALLI 408  
 QY 144 VVGDSLAMDVSV-----HHNSTLRYSVSLILGYGYGDI IKDSEKRWLGLARYDF 195  
 DB 409 ILGRISLDIAQVVRWKSPPSAEVSPTVRYAASFAGYGYGEVIRESEKYRMWGPARYDF 468  
 QY 196 SGLKTFLSHHCYEGTVSFLPAOHT-----VGSPRDRKP-CRAGCFVC 236  
 DB 469 SGTWFLKHSYEAKVAFLENGVTHSLTASANNANGVQTLOYHQRIRKTYCRNCLIC 528  
 QY 237 R---QSKQLEBEOKKALYGLEAAEDVEWQVVCCKFLAINATNMSACRRSPRGLSPAA 293  
 DB 529 KGTSTSEQNSENEDPDS---SRTACETPKWWSKGRFLSVGAAVISCRNERAPDGLVADA 585  
 QY 294 HLGSSDLILIRKSRNFLRLIRHTNQDQ 326  
 DB 586 HLSDGFLHLLLRDCP-----LPFYLWHLTQ 611  
 RESULT 14  
 Q6UZF6 PRELIMINARY; PRT; 532 AA.  
 AC Q6UZF6;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DE Ceramide kinase-like protein.

GN Name=CERKL;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=14681825;  
 RA Tuson M., Marfany G., Gonzalez-Duarte R.;  
 RT "Mutation of CERKL, a novel human ceramide kinase gene, causes  
 RT autosomal recessive retinitis pigmentosa (RP26).";  
 RL Am. J. Hum. Genet. 74:128-138(2004).  
 DR EMBL; AY357073; AAR13670.1; --  
 DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.  
 DR GO; GO:0007205; P:protein kinase C activation; IEA.  
 DR InterPro; IPR001206; DAGKC.  
 DR Pfam; PF00781; DAGK\_cat; 1.  
 DR ProDom; PD005043; DAGKC; 1.  
 KW Kinase.  
 SQ SEQUENCE 532 AA; 59602 MW; C7E590F7C25EED1 CRC64;  
 Query Match 25.4%; Score 436; DB 2; Length 532;  
 Best Local Similarity 31.3%; Pred. No. 2.2e-29;  
 Matches 103; Conservative 63; Mismatches 121; Indels 42; Gaps 6;  
 QY 1 PKHLVFINPFGKGQKRIYERKVAFLTLASITTDIIGNKFFYNYVEVITEHANQAKE 60  
 DB 166 PKSLKILNPQSHKKEATQVYKVEPLLLKLAGIKTDV-----TIMEYEGHALS 214  
 QY 61 TLVEINIDKYDGIYCVGGDMFSEVLHGLIGRTORSAGVDQNHPRALVLP--SSLRIGII 118  
 DB 215 LKSCCEQGFDPGVVVCVGGDGSSEVAHALLRAOKNAGMETDR---ILTPVRAQLPLGLI 271  
 QY 119 PAGSTDCVYSTVGTSDAETSALHI VVGDSLAMDVSVVHNSSTLIRYSVSLILGYGYGDI 178  
 DB 272 PAGSTNVLAHSLHGVPHVITATLH INGHVQLVDVCTFSTAGKLLRFGFSAM-FGFGRT 330  
 QY 179 IKDSEKRWLG-LARYDPSGLKTFLSHHCYEGTVSFLPAOHTVGSVPRDRKPCRAGCFVC 237  
 DB 331 LALAEKYRMWSPNRRDFAVVKALAKLKAEDCEISFLPFNSS----- 372  
 QY 238 QSKQLEBEOKKALYGLEAAEDVEWQVVCCKFLAINATNMSACRRSPRGLSPAHLGD 297  
 DB 373 -----DDVQRRAQGGPKSDCNDQMIOQOFLNVSIMAIPLCSVAPRGLAPNRLNN 426  
 QY 298 GSSDLILIRKSRNFLRLIRHTNQDQ 326  
 DB 427 GSMALLIARTSPEFIFKHLKRYASVKNQ 455  
 RESULT 15  
 Q8L7L1 PRELIMINARY; PRT; 485 AA.  
 AC Q8L7L1;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 05-JUN-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Hypothetical protein At4g21540.  
 GN Name=At4g21540;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Tripp M., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,  
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,  
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,  
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,  
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,  
 RA Ecker J., Theologis A., Davis R.W.;  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2005, 15:12:23 ; Search time 7333.65 Seconds  
(without alignments)  
10664.096 Million cell updates/sec

Title: US-10-631-958-9  
Perfect score: 1614  
Sequence: 1 atggggggcagcggggggcgc.....agccagactcacagctga 1614

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_ats.\*  
12: gb\_ey.\*  
13: gb\_un.\*  
14: gb\_vl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1614	100.0	1654	9	CR456404 Homo sapi
2	1614	100.0	2042	9	AJ457828 Homo sapi
3	1614	100.0	4413	6	AX457006 Sequence
4	1612.4	99.9	4432	6	AR541900 Sequence
5	1612.4	99.9	4445	9	AB079066 Homo sapi
6	1607.6	99.6	4463	6	BD102675 Ceramide
7	1447	89.7	4171	6	BD183468 Novel gen
8	1447	89.7	4171	9	AB051433 Homo sapi
9	1446.4	89.6	1840	6	AX224383 Sequence
10	1444	89.5	1772	9	BC067255 Homo sapi
11	1209.8	75.0	1459	6	CQ730476 Sequence
12	1120	69.4	2830	10	AB079067 Mus muscu
13	903	55.9	979	6	AX456998 Sequence
14	856.2	53.0	3661	10	AX129416 Mus muscu
15	656.8	40.7	1450	5	CR386590 Gallus ga
16	593.8	36.8	2494	5	BC074350 Xenopus 1
17	512.4	31.7	1520	5	BC074110 Xenopus 1
18	475	29.4	550	6	AX457005 Sequence
19	464.4	28.8	474	6	AX457001 Sequence

c	20	320.2	19.8	329	6	AX457002	Sequence
	21	312	19.3	522	6	AX224385	Sequence
	22	298	18.5	3039	9	BC004278	Homo sapi
	23	298	18.5	3074	9	HSB04105	Homo sapi
	24	243.6	15.1	564	6	AX718996	Sequence
c	25	211	13.1	70890	9	HSDA59H18	Human DNA
	26	211	13.1	181772	2	AC007689	Homo sapi
c	27	170.4	10.6	240483	2	AC095955	Rattus no
c	28	170.4	10.6	260331	2	AC135400	Rattus no
	29	167.2	10.4	219077	2	AC116764	Mus muscu
	30	167	10.3	167	6	AX457003	Sequence
	31	153	9.5	153	6	AX457004	Sequence
	32	150	9.3	382	6	AX224387	Sequence
c	33	148.6	9.2	181772	2	AC007689	Homo sapi
c	34	144.2	8.9	89319	9	HSBK29F11	Human DNA
	35	126.6	7.8	1323	5	CR760136	Xenopus t
	36	124.6	7.7	2333	3	AK112750	Ciona int
	37	108.6	6.7	153756	5	BX248240	Zebrafish
c	38	106.2	6.6	325530	2	AC125699	Rattus no
c	39	100.6	6.2	1084	6	AR510195	Sequence
c	40	100.6	6.2	3178	3	BT014925	Drosophil
	41	99	6.1	1810	3	AY061001	Drosophil
	42	99	6.1	2064	6	AR509346	Sequence
	43	99	6.1	2173	6	CQ608040	Sequence
	44	99	6.1	10337	6	CQ608039	Sequence
c	45	99	6.1	22353	2	AC017269	Drosophil

ALIGNMENTS

RESULT 1  
CR456404  
LOCUS  
DEFINITION Homo sapiens BK29F11.1 full length open reading frame (ORF) CDNA  
CR456404 1654 bp mRNA linear PRI 25-MAY-2004  
clone (CDNA clone C22ORF:PGEM.BK29F11.1).  
ACCESSION CR456404.1 GI:47678338  
VERSION CR456404.1  
KEYWORDS CDNA; chromosome 22; ORF.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1654)  
Collins, J.E., Wright, C.L., Edwards, C.A., Davis, M.P., Grinham, J.A.,  
Cole, C.G., Goward, M.E., Aguado, B., Malliya, M., Moktab, Y.,  
Huckle, E.J., Beare, D.M. and Dunham, I.  
Direct Submission  
Submitted (24-MAY-2004) Sanger Institute, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: c22@sanger.ac.uk Manuscript  
Sanger Institute name : PGEM.BK29F11.1  
Homo sapiens CDNA sequence. This sequence was generated as part of  
The Wellcome Trust Sanger Institute program to isolate cDNA clones  
representing the full length open reading frame of well annotated  
protein coding genes on human chromosomes 22. For more information  
see http://www.sanger.ac.uk/HGP/Chr22/.

FEATURES

source  
1. 1654  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="22"  
/clones="PGEM.BK29F11.1"  
/lab\_host="JM109"  
1. 1654  
/gene="BK29F11.1"  
11. 1624  
/gene="BK29F11.1"  
/codon\_start=1  
/protein\_id="CAG30290.1"  
/db\_xref="GI:47678339"  
gene  
1. 1654  
/gene="BK29F11.1"  
CDS  
11. 1624  
/gene="BK29F11.1"  
/codon\_start=1  
/protein\_id="CAG30290.1"  
/db\_xref="GI:47678339"

PGADACSPVPVSEIIAIVETDVHGKHQSGKQMKPYAFTVHCVKRRHRHWQAQV

TFWCPBEOCLHWLQTLREMLEKLTSTRPKHLLVFINPFGKGQGR1YERKVAFLFTL  
ASHTDIIVTSHANQAKETLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVD  
ONHPRALVPSLRIIGIIPAGSTDCVYSTVTGSDAETSAHLHVVGDLSAMDVSSVHH  
NSTLLRYSVLGYPFGVD11KDEKRWLGLARYDSGLKTFLSHHCYEGTWSFLPA  
QHTVSPDRXPCRAGFVCRSQQLKEEOKKALYGLEAABEDVEEHQVVCGRFLAIN  
ATNWSCARRSPRGLSPAHHUGDSSDLILIRKSRNFURFLIRHNTQDDPFTFV  
EYVRVKFQFTSKMEDESDLEKGGKRFHICSSHPSCCCTVSNSSWNCDEVLHS  
PAIEVRVHCQLVRLFARGIEBNPKPDHS"

ORIGIN

Query Match	100.0%;	Score 1614;	DB 9;	Length 1654;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1614;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGGGGCGAGCGGGCGCGGAGCGCTGCAATCCGTGCTGTGGGTGAAGCAGCAGCGC	60	
Db	11	ATGGGGCGAGCGGGCGCGGAGCGCGCTGCAATCCGTGCTGTGGGTGAAGCAGCAGCGC	70	
Qy	61	TGCGCGGTGAGCTGGAGCCCGCGGCGCTCTGCTGCGCTGGTGGCGAGCCCGGGGCC	120	
Db	71	TGCGCGGTGAGCTGGAGCCCGCGGCGCTCTGCTGCGCTGGTGGCGAGCCCGGGGCC	130	
Qy	121	GGAGCGGCGCCCGCGGCGGAGTGCCTGCTCTGTGCTGTGATCTGAGATCATCGCGTT	180	
Db	131	GGAGCGGCGCCCGCGGCGGAGTGCCTGCTCTGTGCTGTGATCTGAGATCATCGCGTT	190	
Qy	191	GAGGAAACAGACGTTTACGGGAAACATCAAGCGAGTGGAAATGCGAGAAATGGAAG	240	
Db	191	GAGGAAACAGACGTTTACGGGAAACATCAAGCGAGTGGAAATGCGAGAAATGGAAG	250	
Qy	241	CCTTACGCTTTTACAGTTTCACTGTGTAAGAGAGACGACGCGCACCGCTGGAAGTGGCG	300	
Db	251	CCTTACGCTTTTACAGTTTCACTGTGTAAGAGAGACGACGCGCACCGCTGGAAGTGGCG	310	
Qy	301	CAGTGACTTCTGTGTGTCAGAGAGACGCTGTGTCTGCTGTGCTGCGACGACCTGCGG	360	
Db	311	CAGTGACTTCTGTGTGTCAGAGAGACGCTGTGTCTGCTGTGCTGCGACGACCTGCGG	370	
Qy	361	GAGATGCTGGAGAGCTGACGCTCAGACCAAGCAATTTACTGTTATTTATCAACCCGTTT	420	
Db	371	GAGATGCTGGAGAGCTGACGCTCAGACCAAGCAATTTACTGTTATTTATCAACCCGTTT	430	
Qy	421	GGAGAAAGGACGAGGCAAGCGGATATATGAAAGAAAGTGGACCACTGTTCAACCTTA	480	
Db	431	GGAGAAAGGACGAGGCAAGCGGATATATGAAAGAAAGTGGACCACTGTTCAACCTTA	490	
Qy	481	GCCTCCATCAGCATGATCATGCTGTACTGAAACATGCTATCAGGCGCAAGGAGACTGTG	540	
Db	491	GCCTCCATCAGCATGATCATGCTGTACTGAAACATGCTATCAGGCGCAAGGAGACTGTG	550	
Qy	541	TATGAGATTAAATAGACAAATACAGCGCATCGTCTGTGTCGGCGGAGATGTTATGTC	600	
Db	551	TATGAGATTAAATAGACAAATACAGCGCATCGTCTGTGTCGGCGGAGATGTTATGTC	610	
Qy	601	AGCGAGGTGCTGACGCTGCTGATTGGGAGGACGACGAGAGCGCGGGGTGCAACGAGAC	660	
Db	611	AGCGAGGTGCTGACGCTGCTGATTGGGAGGACGACGAGAGCGCGGGGTGCAACGAGAC	670	
Qy	661	CACCCCGGGCTGTGCTGGTCCAGTAGCTCCGGATTGGAATCATTTCCCGGAGGTCAT	720	
Db	671	CACCCCGGGCTGTGCTGGTCCCGGATGAGCTCCGGATTGGAATCATTTCCCGGAGGTCAT	730	
Qy	721	ACGGAATCGCTGTGTTACTTCAACCGTGGGACGACGAGCAACCTCGGCGTGCAT	780	
Db	731	ACGGAATCGCTGTGTTACTTCAACCGTGGGACGACGAGCAACCTCGGCGTGCAT	790	
Qy	781	ATCGTTGTTGGGACTCGCTGCGCATGATGTCTCTCAGTCCACCAACAGCAGCACTC	840	
Db	791	ATCGTTGTTGGGACTCGCTGCGCATGATGTCTCTCAGTCCACCAACAGCAGCACTC	850	
Qy	841	CTTGCTACTCGGTGCTGCTGGCTACCGCTTCTACGGGACATCATCAAGGACAGT	900	
Db	851	CTTGCTACTCGGTGCTGCTGGCTACCGCTTCTACGGGACATCATCAAGGACAGT	910	

Qy	901	GAGAAGAAACGGTGGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACCTTCCTC	960	
Db	911	GAGAAGAAACGGTGGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACCTTCCTC	970	
Qy	961	TCCACCACTGCTATGAAAGGAGACAGTGTCTTCTCTCCCTGCAACACACAGGTTGGATCT	1020	
Db	971	TCCACCACTGCTATGAAAGGAGACAGTGTCTTCTCTCCCTGCAACACACAGGTTGGATCT	1030	
Qy	1021	CCAGGGATAGGAAGCCCTCCCGGCGAGGATGCTTTGTTTGCAGGCAAAACAGCAGCAG	1080	
Db	1031	CCAGGGATAGGAAGCCCTCCCGGCGAGGATGCTTTGTTTGCAGGCAAAACAGCAGCAG	1090	
Qy	1081	CTGAGAGGAGCAGAGAAAGACACTGTATGGTTTGAAGCTGCGGAGGACGTTGGAGGAG	1140	
Db	1091	CTGAGAGGAGCAGAGAAAGACACTGTATGGTTTGAAGCTGCGGAGGACGTTGGAGGAG	1150	
Qy	1141	TGGCAATCTGCTGTGGGAAAGTTTCTGGCCATCAATGCCCAACACATGTCTGTGCTTGT	1200	
Db	1151	TGGCAATCTGCTGTGGGAAAGTTTCTGGCCATCAATGCCCAACACATGTCTGTGCTTGT	1210	
Qy	1201	CGCGGAGCCCGAGGGGCGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCTGACCTC	1260	
Db	1211	CGCGGAGCCCGAGGGGCGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCTGACCTC	1270	
Qy	1261	ATCTCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTTCTCATCAGGCAACCAAC	1320	
Db	1271	ATCTCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTTCTCATCAGGCAACCAAC	1330	
Qy	1321	CAGCAGACCAAGTTGACTTCACTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTT	1380	
Db	1331	CAGCAGACCAAGTTGACTTCACTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTT	1390	
Qy	1381	ACGTCGAAGCAGATGGAGGATGAGCAGCGACTCAAGGAGGGGGGAGAGCGCTTT	1440	
Db	1391	ACGTCGAAGCAGATGGAGGATGAGCAGCGACTCAAGGAGGGGGGAGAGCGCTTT	1450	
Qy	1441	GGGCACTTTTGCAGCAGCCACCCCTCTGCTGTGCTGCAACCGTCTCCAACAGCTCCTGGAAC	1500	
Db	1451	GGGCACTTTTGCAGCAGCCACCCCTCTGCTGTGCTGCAACCGTCTCCAACAGCTCCTGGAAC	1510	
Qy	1501	TGCGAGCGGAGGTCCTGCGACAGCCCTGCGCATGAGGTGAGTCCAGTCCAGTCCGCTGTT	1560	
Db	1511	TGCGAGCGGAGGTCCTGCGACAGCCCTGCGCATGAGGTGAGTCCAGTCCAGTCCGCTGTT	1570	
Qy	1561	CGACTCTTTCAGAGGAAATTGAAGAGAATCCGAAGCCAGACTCACAACAGCTGA	1614	
Db	1571	CGACTCTTTCAGAGGAAATTGAAGAGAATCCGAAGCCAGACTCACAACAGCTGA	1624	

RESULT 2

HSA457828  
LOCUS HSA457828 2042 bp mRNA linear PRI 19-APR-2002  
DEFINITION Homo sapiens mRNA for putative lipid kinase (LK4 gene).  
ACCESSION AJ457828  
VERSION AJ457828.1 GI:20269072

KEYWORDS lipid kinase; LK4 gene.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 Van Veldhoven, P.P.  
AUTHORS A search for lipid kinases  
TITLE Unpublished  
JOURNAL (bases 1 to 2042)  
REFERENCE 2 Van Veldhoven, P.P.  
AUTHORS Direct Submission  
TITLE Submitted (18-APR-2002) Van Veldhoven P.P., Farmakologie,  
JOURNAL K.U.Leuven, Herestraat, B-3000 Leuven, BELGIUM

FEATURES

source  
1. .2042  
/organism="Homo sapiens"

```

/mol_type="mrna"
/db_xref="taxon:9606"
/chromosome="22"
/map="22q13"
1..2042
/gene="LK4"
71..1684
/gene="LK4"
/codon_start=1
/product="putative lipid kinase"
/protein_id="CAD2984.1"
/db_xref="GI:20269073"
/db_xref="GOA:O8TCT0"
/db_xref="UniProt/Swiss-Prot:O8TCT0"
/translation="MGATGAEPLOSIVLWVKQORCAVSLRPARALLRWRSPGPAG
PADACSVPSVSIIVAEETDVHKGQSGKQMKPYAFTVCHVKRARRHRWKAQV
TPWCEPOLQLMQLTREMELFKLSRPKHLIVFNPFGKGQKRIYERKVAPLFTL
ASITDIIVTEHANAKETLYEINDIKVDYICVCGDGMFSEVLHGLIGRTORSAGVD
QNHPRALVPSSLRIGIIPAGSTDCVYSTVGTSDAETSALHIIVGDSLMDVSVVH
NSTLARYSVLLGIFYGDIINDSEKRWGLIARYDFSLKTFILSHHCIEGTVSPLPA
QHTVSPDRKFCRAGCFVCRQSKQLEBEQKALYLEAAEDVEEWVVCQKFLAIN
ATNMSCARRSPRGLSPAHLGDSGLILIRKSRFNFLRLHRTNQDFV
EYRVKFTQFTSKHMEDESDLKEGKKRFGHICSSHPSCCTVSNSSWNCDEVLHS
PAIEVRVHQLVRLPARGIEENKPDHS"
```

ORIGIN

```

Query Match      100.0%; Score 1614; DB 9; Length 2042;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1614; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ATGGGGCGACGGGGCGGCGCGCGTGCATCCGTGCTGTGGTGAAGCAGCAGCGC 60
DB      71  ATGGGGCGACGGGGCGGCGCGCGTGCATCCGTGCTGTGGTGAAGCAGCAGCGC 130

QY      61  TCGCGCGTGAGCTGAGCGCGCGCGGCTCTGCTCGCTGTGGCGGAGCCCGGGGCC 120
DB      131  TCGCGCGTGAGCTGAGCGCGCGCGGCTCTGCTCGCTGTGGCGGAGCCCGGGGCC 190

QY      121  GGAGCGGCGCGCGCGCGCGAGTGCCTGCTGTGCTGTGCTGTGATCTGAGATCATGCCGTT 180
DB      191  GGAGCGGCGCGCGCGCGCGAGTGCCTGCTGTGCTGTGATCTGAGATCATGCCGTT 250

QY      181  GAGGAAACAGAGCTTACAGTTCACGTGTAAAGAGAGCAGCAGCGCACCGCTGGAAGTGGCG 300
DB      251  GAGGAAACAGAGCTTACAGTTCACGTGTAAAGAGAGCAGCAGCGCACCGCTGGAAGTGGCG 310

QY      241  CCTTACGCTTTTACAGTTCACGTGTAAAGAGAGCAGCAGCGCACCGCTGGAAGTGGCG 300
DB      311  CCTTACGCTTTTACAGTTCACGTGTAAAGAGAGCAGCAGCGCACCGCTGGAAGTGGCG 370

QY      301  CAGGTGACTTTCTGTGTCCAGAGGAGCAGTGTCTCACTTGTGTGCTGAGACCCCTGCGG 360
DB      371  CAGGTGACTTTCTGTGTCCAGAGGAGCAGTGTCTCACTTGTGTGCTGAGACCCCTGCGG 430

QY      361  GAGATGCTGGAGAGTGCAGTGCAGACCAAGCAATTTACTGGTATTTATCAACCCGTTT 420
DB      431  GAGATGCTGGAGAGTGCAGTGCAGACCAAGCAATTTACTGGTATTTATCAACCCGTTT 490

QY      421  GGAGAAAGGACAAGGCGGATATATGAAAGAAAGTGGCACCACTGTTTCACTTCA 480
DB      491  GGAGAAAGGACAAGGCGGATATATGAAAGAAAGTGGCACCACTGTTTCACTTCA 550

QY      481  GCCTCATCAACCTGATCATCTGTTACTGAAACATGCTTAATCAGGCCAAGGAGACTCTG 540
DB      551  GCCTCATCAACCTGATCATCTGTTACTGAAACATGCTTAATCAGGCCAAGGAGACTCTG 610

QY      541  TATGAGTTTAAATAGACAAATACAGCGCATCGTCTGTGTCGGCGGAGATGTTATGTT 600
DB      611  TATGAGTTTAAATAGACAAATACAGCGCATCGTCTGTGTCGGCGGAGATGTTATGTT 670

QY      601  ACCGAGGTGCTCAGCGTCTGATTCGGGAGGAGCAGAGGAGCGCGCGGCTCGACCAAGAC 660
DB      671  ACCGAGGTGCTCAGCGTCTGATTCGGGAGGAGCAGAGGAGCGCGCGGCTCGACCAAGAC 730
```

```

QY      661  CACCCCGGGCTGTGCTGCTCCCGAGTAGCCTCCGGATTGGAATCATTTCCCGCAGGTCA 720
DB      731  CACCCCGGGCTGTGCTGCTCCCGAGTAGCCTCCGGATTGGAATCATTTCCCGCAGGTCA 790

QY      721  ACGGACTCGGTGTGTACTCCACCGTGGGACACGAGCAGCAGAAACCTCGCGCTGCTAT 780
DB      791  ACGGACTCGGTGTGTACTCCACCGTGGGACACGAGCAGCAGAAACCTCGCGCTGCTAT 850

QY      781  ATCGTGTGTGGGACTCGCTGGCCATGATGTGCTCTCAGTCCACCAACAGACACATCTC 840
DB      851  ATCGTGTGTGGGACTCGCTGGCCATGATGTGCTCTCAGTCCACCAACAGACACATCTC 910

QY      841  CTTGCTACTCGGTGTGCTGCTGGGCTACCGGCTTCTAAGGGGACATCATCAAGACAGT 900
DB      911  CTTGCTACTCGGTGTGCTGCTGGGCTACCGGCTTCTAAGGGGACATCATCAAGACAGT 970

QY      901  GAGAAAGAACGCTGTGGGCTCTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCT 960
DB      971  GAGAAAGAACGCTGTGGGCTCTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCT 1030

QY      961  TCCACCACTGCTATGAAGGAGCAGTGTCTTCTCTCCCTGCAACACACACGCTGGGATCT 1020
DB      1031  TCCACCACTGCTATGAAGGAGCAGTGTCTTCTCTCCCTGCAACACACACGCTGGGATCT 1090

QY      1021  CCAAGGATAGAAAGCCCTGCGGGCAGGATGCTTTGTTTGGAGGAAAGCAAGCAGCAG 1080
DB      1091  CCAAGGATAGAAAGCCCTGCGGGCAGGATGCTTTGTTTGGAGGAAAGCAAGCAGCAG 1150

QY      1081  CTGAGGAGGAGCAGAAAGAACGACTGTATGTTTGAAGCTGCGGAGCAGCTGGAGGAG 1140
DB      1151  CTGAGGAGGAGCAGAAAGAACGACTGTATGTTTGAAGCTGCGGAGCAGCTGGAGGAG 1210

QY      1141  TGGCAAGTCTGTGTGGGAAGTTTCTGCCATCAATGCAACAAACATGCTCTGCTGTTGT 1200
DB      1211  TGGCAAGTCTGTGTGGGAAGTTTCTGCCATCAATGCAACAAACATGCTCTGCTGTTGT 1270

QY      1201  CGCGGAGCCCGAGGGGCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTCTGACCTC 1260
DB      1271  CGCGGAGCCCGAGGGGCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTCTGACCTC 1330

QY      1261  ATCTCTATCCGGAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGCAGACCAAC 1320
DB      1331  ATCTCTATCCGGAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGCAGACCAAC 1390

QY      1321  CAGCAGGACCAAGTTTGAATTTGTTGAAAGTTTATCGCTCAAGAAATTCAGATTT 1380
DB      1391  CAGCAGGACCAAGTTTGAATTTGTTGAAAGTTTATCGCTCAAGAAATTCAGATTT 1450

QY      1381  ACCTCGAAGCAGATGAGGATGAGGACAGCAGCCTCAAGAGGGGGGAGAAAGCGCTTT 1440
DB      1451  ACCTCGAAGCAGATGAGGATGAGGACAGCAGCCTCAAGAGGGGGGAGAAAGCGCTTT 1510

QY      1441  GGGCAGATTTGACAGCAGCCCTCTGCTGCTGCAACCGCTTCCACAGCTCTCTGGAAC 1500
DB      1511  GGGCAGATTTGACAGCAGCCCTCTGCTGCTGCAACCGCTTCCACAGCTCTCTGGAAC 1570

QY      1501  TGGCAGGGGAGGCTCTGCAAGCCCTGCCATCGAGGTTCAGAGTCCACTGCGCAGCTGTT 1560
DB      1571  TGGCAGGGGAGGCTCTGCAAGCCCTGCCATCGAGGTTCAGAGTCCACTGCGCAGCTGTT 1630

QY      1561  CGACTTTTTCAGAGGAAATGAGAGAAATCCGAAGCCAGACTCACACAGCTGA 1614
DB      1631  CGACTTTTTCAGAGGAAATGAGAGAAATCCGAAGCCAGACTCACACAGCTGA 1684
```

RESULT 3  
AX457006  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS

AX457006  
Sequence 9 from Patent WO0228906.  
AX457006  
AX457006.1  
GI:21715795

AX457006  
4413 bp  
DNA  
linear  
PAT 06-JUL-2002

SOURCE	Homo sapiens (human)	Db	616	TATGAGATTAACTAGACAAAATACGACGGCATCGTGTGTGTCGGCGGAGATGGTATGTC	675
ORGANISM	Homo sapiens	Qy	601	AGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGGAGCGCGGGGTTCGACCAAGAAC	660
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Db	676	AGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGGAGCGCGGGGTTCGACCAAGAAC	735
AUTHORS	Kossida, S. and Encinas, J.	Qy	661	CACCCCGGGGTGCTGTGGTCCCGAGTAGCTCCGGATTTGGAATCATTTCCCGCAGGGTCA	720
TITLE	Regulation of human sphingosine kinase-like protein	Db	736	CACCCCGGGGTGCTGTGGTCCCGAGTAGCTCCGGATTTGGAATCATTTCCCGCAGGGTCA	795
JOURNAL	Patent: WO 0228906-A 9 11-APR-2002;	Qy	721	ACGAGCTCGGTGTGTTACTCCACCGTGGGACACAGCAGCAGAGAAAACCTCGGCGCTGGAT	780
Patent: WO 0228906-A 9 11-APR-2002;		Db	796	ACGAGCTCGGTGTGTTACTCCACCGTGGGACACAGCAGCAGAGAAAACCTCGGCGCTGGAT	855
Bayer Aktiengesellschaft (DE)		Qy	781	ATCGTGTGTTGGGAGTCTGCTGGCCATGGATGTGCTCAGTCCACCAACACAGCACACTC	840
Location/Qualifiers		Db	856	ATCGTGTGTTGGGAGTCTGCTGGCCATGGATGTGCTCAGTCCACCAACACAGCACACTC	915
1..4413		Qy	841	CTTCGCTACTCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAGGACAGT	900
/organism="Homo sapiens"		Db	916	CTTCGCTACTCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAGGACAGT	975
/mol_type="unassigned DNA"		Qy	901	GAGAAAGAACCGTGTGGTCTGTCAGATACGACTTTTTCAGGTTTAAAGACCTTCCTC	960
/db_xref="taxon:9606"		Db	976	GAGAAAGAACCGTGTGGTCTGTCAGATACGACTTTTTCAGGTTTAAAGACCTTCCTC	1035
76..1689		Qy	961	TCCACCACTGCTATGAGGAGACAGTGTCTTCTCCTGCACAAACACAGCGTGGGATCT	1020
/note="unnamed protein product"		Db	1036	TCCACCACTGCTATGAGGAGACAGTGTCTTCTCCTGCACAAACACAGCGTGGGATCT	1095
/codon_start=1		Qy	1021	CCAGGGATAGGAAGCCCTGCCGGGACAGTGTCTTGTGTGAGGCAAGCAAGCAGCAG	1080
/protein_id="CAD38256.1"		Db	1096	CCAGGGATAGGAAGCCCTGCCGGGACAGTGTCTTGTGTGAGGCAAGCAAGCAGCAG	1155
/db_xref="GI:21715796"		Qy	1081	CTGAGGAGGAGCAGAAAGACACTGTATGCTTTGGAAGCTGCGGAGGAGCTGGAGGAG	1140
/translation="MGATGAAPLQSVLWVKQORCAVSLPARALLRWRSPPGAGAPGADACVPVSEIITAVEEDVHGKQSGKWKQEKYAFVTHVCVKRABRHWKQAVTFWCPERQCLHLWQLTREMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLQNHPRVLVPSLRIGTIPAGSTDVCYSTVGTSDAETSAHLHVVGDSLAMDVSVFHNSTLLRYSVLLGTFYGDIIKDSKKRWGLGARYDFSLKTFLSHHCYEGTGVFLPAQHTVSPDRKPCRCAGCFVCRSQKQLEEKALYGLEAEDVEEVQVCGKFLFNAATNWSACRRSPRGLSPAHLGDGSDLLILRKSRNFRLFRLRHTNQDQDPFTFVEYRVKQFTFSKMEDESDLKGGKKRFGHICSSHPSCCTVSNSSWNCDEGVILHSPAIERVHCQLVRLFARGIENPKPDHS"		Db	1156	CTGAGGAGGAGCAGAAAGACACTGTATGCTTTGGAAGCTGCGGAGGAGCTGGAGGAG	1215
		Qy	1141	TGGCAAGTCGCTCTGGGAAAGTTTCTGCCATCAATGCACAAACATGCTCTGCTGTGT	1200
		Db	1216	TGGCAAGTCGCTCTGGGAAAGTTTCTGCCATCAATGCCAACAAACATGCTCTGCTGTGT	1275
		Qy	1201	CGCCGAGCCCGCCAGGGGCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTTGACCTC	1260
		Db	1276	CGCCGAGCCCGCCAGGGGCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTTGACCTC	1335
		Qy	1261	ATCCTCATTCGGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGCAGCAGCAAC	1320
		Db	1336	ATCCTCATTCGGAAATGCTCCAGGTTCAATTTTCTGAGATTTTCTCATCAGCAGCAGCAAC	1395
		Qy	1321	CAGCAGGACCAAGTTTGACTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTT	1380
		Db	1396	CAGCAGGACCAAGTTTGACTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTT	1455
		Qy	1381	ACGTTCGAAGCAGCATGGAGGATGAGACAGCGACCTCAAGGAGGGGGGAGAAAGCGCTTT	1440
		Db	1456	ACGTTCGAAGCAGCATGGAGGATGAGACAGCGACCTCAAGGAGGGGGGAGAAAGCGCTTT	1515
		Qy	1441	GGGACATTTGACAGCAGCAGCCCTCTGCTGTGTCACCGTCTCAACAGCTCTCGAAC	1500
		Db	1516	GGGACATTTGACAGCAGCAGCCCTCTGCTGTGTCACCGTCTCAACAGCTCTCGAAC	1575
		Qy	1501	TGCCAGCGGGAGGTCTTGCAACAGCCCTGCCATCAGAGGTTCAGAGTCCACTGCCAGCTGGT	1560
		Db	1576	TGCCAGCGGGAGGTCTTGCAACAGCCCTGCCATCAGAGGTTCAGAGTCCACTGCCAGCTGGT	1635
		Qy	1561	CGACTCTTTGACAGGAAATTGAAAGAGAAATCCGAAGCCAGACTCAGCAGCTGA	1614
		Db	1636	CGACTCTTTGACAGGAAATTGAAAGAGAAATCCGAAGCCAGACTCAGCAGCTGA	1689
ORIGIN					
Query Match	100.0%; Score 1614; DB 6; Length 4413;				
Best Local Similarity	100.0%; Pred. No. 0;				
Matches 1614; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1	ATGGGGGCGACGGGGCGGAGCGCGTCAATCCGTCGTGTGGTGAAGCAGCAGCGC	60		
Db	76	ATGGGGGCGACGGGGCGGAGCGCGTCAATCCGTCGTGTGGTGAAGCAGCAGCGC	135		
Qy	61	TGCGCGGTGAGCTGGAGCCCGCGGCTCTGCGCTGGTGGCGAGCCCGGGGCC	120		
Db	136	TGCGCGGTGAGCTGGAGCCCGCGGCTCTGCGCTGGTGGCGAGCCCGGGGCC	195		
Qy	121	GGAGCCGCGCCCGCGCGGATCGCTGCTGTGCTGTATCTGAGATCATCGCGTT	180		
Db	196	GGAGCCGCGCCCGCGCGGATCGCTGCTGTGCTGTATCTGAGATCATCGCGTT	255		
Qy	181	GAGGAAACAGAGCTTACCGGAAACATCAAGGACAGTGGAAATGCGAAATGGAAG	240		
Db	256	GAGGAAACAGAGCTTACCGGAAACATCAAGGACAGTGGAAATGCGAAATGGAAG	315		
Qy	241	CTTTACGCTTTTACAGTTTACCTGTGTAAGAGAGCAGCAGCCACCGTGAAGTGGCG	300		
Db	316	CTTTACGCTTTTACAGTTTACCTGTGTAAGAGAGCAGCAGCCACCGTGAAGTGGCG	375		
Qy	301	CAGGTGACTTTCTGTGTCCAGAGAGCAGCTGTGCTGCTGCTGAGACCTCGCG	360		
Db	376	CAGGTGACTTTCTGTGTCCAGAGAGCAGCTGTGCTGCTGCTGAGACCTCGCG	435		
Qy	361	GAGATGCTGGAGAGCTGACGTCCAGACCAAGCATTTTCTGGTATTTATCAACCCGTTT	420		
Db	436	GAGATGCTGGAGAGCTGACGTCCAGACCAAGCATTTTCTGGTATTTATCAACCCGTTT	495		
Qy	421	GGAGAAAGGACCAAGCGGATATGAAAGAAAGTGGCCACTGTTTCACTTTA	480		
Db	496	GGAGAAAGGACCAAGCGGATATGAAAGAAAGTGGCCACTGTTTCACTTTA	555		
Qy	481	GCCTCCATCACCAGTACATCATCTGTTACTGAAATGCTTAATCAGGCAAGAGAGCTCTG	540		
Db	556	GCCTCCATCACCAGTACATCATCTGTTACTGAAATGCTTAATCAGGCAAGAGAGCTCTG	615		
Qy	541	TATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGGTATGTC	600		

LOCUS AR541900 4432 bp DNA linear PAT 08-OCT-2004  
DEFINITION Sequence 148 from patent US 6743619.  
ACCESSION AR541900  
VERSION AR541900.1 GI:53933980  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 4432)  
AUTHORS Tang,Y.T., Zhou,P., Goodrich,R., Liu,C., Asundi,V., Ren,F.,  
Zhang,J., Zhao,Q.A., Yang,Y., Xue,A.J., Wehrman,T., Wang,J.-R.,  
Wang,D. and Drmanac,R.T.  
TITLE Nucleic acids and polypeptides  
JOURNAL Patent: US 6743619-A 148 01-JUN-2004;  
FEATURES Location/Qualifiers  
source 1..4432  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 99.9%; Score 1612.4; DB 6; Length 4432;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1613; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGGGCGACGGGGCGCGGAGCGCTGCAATCGTCTGTTGGTGAAGCAGCGC 60  
Db 1204 ATGGGGGCGACGGGGCGCGGAGCGCTGCAATCGTCTGTTGGTGAAGCAGCGC 1263

QY 61 TGGCGCGTGAGCGCTGAGAGCCCGCGGGCTCTGCTGCGCTGTTGGCGAGCCCGGGCC 120  
Db 1264 TGGCGCGTGAGCGCTGAGAGCCCGCGGGCTCTGCTGCGCTGTTGGCGAGCCCGGGCC 1323

QY 121 GAGCGCGCGCCCGCGCGGAGTGCTGCTGTCCTGATCTGAGATCATGCGCGTT 180  
Db 1324 GAGCGCGCGCCCGCGCGGAGTGCTGCTGTCCTGATCTGAGATCATGCGCGTT 1383

QY 181 GAGGAACAGAGCTTACAGGAAACATCAAGCAGTGGAAATGCGAGAAATGGAAG 240  
Db 1384 GAGGAACAGAGCTTACAGGAAACATCAAGCAGTGGAAATGCGAGAAATGGAAG 1443

QY 241 CCTTAGCGCTTTACAGTTCACTGTGTAAAGAGACGACGCGCACCGCTGGAAGTGGCG 300  
Db 1444 CCTTAGCGCTTTACAGTTCACTGTGTAAAGAGACGACGCGCACCGCTGGAAGTGGCG 1503

QY 301 CAGGTGACTTCTGTGTTCCAGAGGAGAGCTGTGTCATTTGGTGTGAGACCCCTGCGG 360  
Db 1504 CAGGTGACTTCTGTGTTCCAGAGGAGAGCTGTGTCATTTGGTGTGAGACCCCTGCGG 1563

QY 361 GAGATGCTGGAGAGCTGACGCTCCAGACCAAGCATTTTACTGGTATTTATCAACCGTTT 420  
Db 1564 GAGATGCTGGAGAGCTGACGCTCCAGACCAAGCATTTTACTGGTATTTATCAACCGTTT 1623

QY 421 GGAGGAAAAGGACAAGCGGCGGATATATGAAAGAAAAGTGGCACCACTGTTCACTTA 480  
Db 1624 GGAGGAAAAGGACAAGCGGCGGATATATGAAAGAAAAGTGGCACCACTGTTCACTTA 1683

QY 481 GCCTCCATCACCACATCATCGTTACTGAACTGTCATGTCATGAGCCCAAGGAGACTCTG 540  
Db 1684 GCCTCCATCACCACATCATCGTTACTGAACTGTCATGTCATGAGCCCAAGGAGACTCTG 1743

QY 541 TATGAGATTAAATAGACAAATACGACGCGCATCGTCTGTGTCGGCGGAGATGGTATGTC 600  
Db 1744 TATGAGATTAAATAGACAAATACGACGCGCATCGTCTGTGTCGGCGGAGATGGTATGTC 1803

QY 601 AGCGAGTGTCTGACCGGTCTGATTTGGGAGGAGCGAGAGAGCGCGGGTCCACAGAAC 660  
Db 1804 AGCGAGTGTCTGACCGGTCTGATTTGGGAGGAGCGAGAGAGCGCGGGTCCACAGAAC 1863

QY 661 CACCCCGCGGCTGTGTCGTCGCCAGTAGCTCCGAGTTGGAATCATTTCCCGAGGGTCA 720  
Db 1864 CACCCCGCGGCTGTGTCGTCGCCAGTAGCTCCGAGTTGGAATCATTTCCCGAGGGTCA 1923

QY 721 ACGGACTGCGTGTGTACTCTCCACCGTGGGACACGAGCGCAGAGAAACCTCGCGCGTGCAT 780

Db 1924 ACGGACTGCGTGTGTGTACTCTCACCGTGGGACACGAGCGAGAAACCTTCGGCGCTGCAT 1983

QY 781 ATCGTTTGTGGGACTCGCTGGCCATGGATGTGTCTCAGTCCACCAACAGCACACTTC 840

Db 1984 ATCGTTTGTGGGACTCGCTGGCCATGGATGTGTCTCAGTCCACCAACAGCACACTTC 2043

QY 841 CTTCCGCTACTCGTGTCTCCTGCTGGGCTACCGGCTTCTACGGGGACATCATCAAGGACAGT 900

Db 2044 CTTCCGCTACTCGTGTCTCCTGCTGGGCTACCGGCTTCTACGGGGACATCATCAAGGACAGT 2103

QY 901 GAGAAGAAAACGGTGTGTGGGCTCTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCCTC 960

Db 2104 GAGAAGAAAACGGTGTGTGGGCTCTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCCTC 2163

QY 961 TCCCAACCATGCTATGAAGGGACAGTGTCTTCTCTCCCTGCAACACACACGCTGGGATCT 1020

Db 2164 TCCCAACCATGCTATGAAGGGACAGTGTCTTCTCTCCCTGCAACACACACGCTGGGATCT 2223

QY 1021 CCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTTGTGGAGCAAGCAAGCAGCAG 1080

Db 2224 CCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTTGTGGAGCAAGCAAGCAGCAG 2283

QY 1081 CTGGAGGAGGAGCACAAGAAAGCACTGTATGTTTGGAGCTGCGGAGACGCTGGAGGAG 1140

Db 2284 CTGGAGGAGGAGCACAAGAAAGCACTGTATGTTTGGAGCTGCGGAGACGCTGGAGGAG 2343

QY 1141 TGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCAACAAACATGCTCTGTGCTTGT 1200

Db 2344 TGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCAACAAACATGCTCTGTGCTTGT 2403

QY 1201 CGCGGAGCCCGCAGGGGCTCTCTCCCGGCTGCCCCAATTGGGAGACGGGTCTTCTGACCTC 1260

Db 2404 CGCGGAGCCCGCAGGGGCTCTCTCCCGGCTGCCCCAATTGGGAGACGGGTCTTCTGACCTC 2463

QY 1261 ATCTCTATCCGAAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCACACCAAC 1320

Db 2464 ATCTCTATCCGAAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCACACCAAC 2523

QY 1321 CAGCAGGACCAAGTTTGAAGTTTATCGCTCAAGAAATTCAGTTT 1380

Db 2524 CAGCAGGACCAAGTTTGAAGTTTATCGCTCAAGAAATTCAGTTT 2583

QY 1381 AGCTCGAAGCACAAGGAGATGAGACAGCAGCCTCAAGAGGGGGGGAAGAGCGCTTT 1440

Db 2584 AGCTCGAAGCACAAGGAGATGAGACAGCAGCCTCAAGAGGGGGGGAAGAGCGCTTT 2643

QY 1441 GGGCACATTTGAGCAGCAGCCCTCTGCTGTCGACCGTCTCAACAGCTCTCGGAAC 1500

Db 2644 GGGCACATTTGAGCAGCAGCCCTCTGCTGTCGACCGTCTCAACAGCTCTCGGAAC 2703

QY 1501 TGCGACGGGAGGTCCTGCACAGCCCTGCCATCGAGGTCCAGGTCCAGTCCAGCTGGTT 1560

Db 2704 TGCGACGGGAGGTCCTGCACAGCCCTGCCATCGAGGTCCAGGTCCAGTCCAGCTGGTT 2763

QY 1561 CGACTCTTTTGCAGGAGAAATTGAAGAAATCGAAGCCAGACTCACAGCTGA 1614

Db 2764 CGACTCTTTTGCAGGAGAAATTGAAGAAATCGAAGCCAGACTCACAGCTGA 2817

RESULT 5  
AB079066  
LOCUS Homo sapiens cerk mRNA for ceramide kinase, complete cds.  
DEFINITION AB079066  
ACCESSION AB079066  
VERSION AB079066.1 GI:21624339  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 Sugiyama,M., Kono,K., Liu,H., Shimizuawa,T., Minekura,H.,





```
QY 1561 CGACTCTTTGGCAGGAGGAAATTGAAGAGAAATCCGAAGCAGACTCACAGCTGA 1614
DB 1684 CGACTCTTTGGCAGGAGGAAATTGAAGAGAAATCCGAAGCAGACTCACAGCTGA 1737

RESULT 6
BD102675          4463 bp          linear          PAT 27-AUG-2002
LOCUS             Ceramide kinase and DNA thereof.
DEFINITION        BD102675
ACCESSION          BD102675
VERSION            BD102675.1 GI:22648249
KEYWORDS           WO 0196575-A/1.
SOURCE             Homo sapiens (human)
ORGANISM           Homo sapiens
REFERENCE           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE              1 (bases 1 to 4463)
JOURNAL            Sugiyura, M., Kono, K. and Kohama, T.
COMMENT            Ceramide kinase and DNA thereof.
                    SANKYO CO LTD, MASAKO SUGIURA, KEITA KONO, TAKAFUMI KOHAMA
                    Patent: WO 0196575-A 1 20-DEC-2001;
                    OS Homo sapiens (human)
                    PN WO 0196575-A/1
                    PD 20-DEC-2001
                    PF 11-JUN-2001 WO 2001JP004889
                    PR 14-JUN-2000 JP OOP 178039
                    PI MASAKO SUGIURA, KEITA KONO, TAKAFUMI KOHAMA
                    PC C12N15/54, C12N9/12, C12N1/21, C12Q1/48, C07K16/40, A61K31/7125, PC
                    A61K48/00,
                    PC A61P3/04, A61P3/10, A61P7/00, A61P25/00, A61P29/00, A61P31/18, PC
                    A61P35/00
                    CC Ceramide kinase and DNA thereof
                    FH Key Location/Qualifiers
                    PT CDS (124)..(1734).
                    Location/Qualifiers
                    1..4463
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"

FEATURES
source

ORIGIN
Query Match 99.6%; Score 1607.6; DB 6; Length 4463;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1610; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGGGGCGACGGGGCGGCGGAGCGCGTGCATCCGTGCTGTGGGTGAAGCAGCAGCGC 60
DB 124 ATGGGGGCGACGGGGCGGCGGAGCGCGTGCATCCGTGCTGTGGGTGAAGCAGCAGCGC 183

QY 61 TCGCCGCTGAGCTTGAAGCCCGCGCGGCTCTGCTGCGCTGTGGCGAGCCCGGGGCC 120
DB 184 TCGCCGCTGAGCTTGAAGCCCGCGCGGCTCTGCTGCGCTGTGGCGAGCCCGGGGCC 243

QY 121 GGAGCCGGCGCCCGGGCGGAGTCCGCTGCTGTGCTGTGCTGTGATCTGAGATCATCGCGTT 180
DB 244 GGAGCCGGCGCCCGGGCGGAGTCCGCTGCTGTGCTGTGCTGTGATCTGAGATCATCGCGTT 303

QY 181 GAGGAAACAGAGCTTTCACGGGAAACATCAAGGACGTGGAAATGGCAGAAATGGAAGA 240
DB 304 GAGGAAACAGAGCTTTCACGGGAAACATCAAGGACGTGGAAATGGCAGAAATGGAAGA 363

QY 241 CTTTACGCTTTTACAGTTTCACTGTGTAAGAGAGCAGCAGCGCACCGCTGGAAGTGGCG 300
DB 364 CTTTACGCTTTTACAGTTTCACTGTGTAAGAGAGCAGCAGCGCACCGCTGGAAGTGGCG 423

QY 301 CAGGTGACTTTCTGTGTGTCAGAGAGCAGTGTGTCACTTGTGCTGCAGACCCCTGCGG 360
DB 424 CAGGTGACTTTCTGTGTGTCAGAGAGCAGTGTGTCACTTGTGCTGCAGACCCCTGCGG 483

QY 361 GAGATGCTGGAGAGCTGACGCTCCAGACCAAGCAATTTACTGGTATTTTCAACCCGTTT 420
DB 484 GAGATGCTGGAGAGCTGACGCTCCAGACCAAGCAATTTACTGGTATTTTCAACCCGTTT 543
```

```
QY 421 GGAGAAAGGACAAAGGCAAGCGGATATATGAAAGAAAAGTGGACCACTGTTTCACTTA 480
DB 544 GGAGAAAGGACAAAGGCAAGCGGATATATGAAAGAAAAGTGGACCACTGTTTCACTTA 603

QY 481 GCCTCCATCACCACCTGACATCATCGTTACTGAAACATGCTAATCAGGCGCAAGGAGACTCTG 540
DB 604 GCCTCCATCACCACCTGACATCATCGTTACTGAAACATGCTAATCAGGCGCAAGGAGACTCTG 663

QY 541 TATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGGTATGTTTC 600
DB 664 TATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGGTATGTTTC 723

QY 601 AGCAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGAGCGCGGGGTGCACAGAAC 660
DB 724 AGCAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGAGCGCGGGGTGCACAGAAC 783

QY 661 CACCCCGGGCTGTGCTGGTCCCGAGTAGCTCCGGAATGGAATCATTTCCCGCAGGGTCA 720
DB 784 CACCCCGGGCTGTGCTGGTCCCGAGTAGCTCCGGAATGGAATCATTTCCCGCAGGGTCA 843

QY 721 ACGGACTGCGTGTGTTACTCCACCGTGGGCAACGACGACGAGAAACCTCGCGCTGCAT 780
DB 844 ACGGACTGCGTGTGTTACTCCACCGTGGGCAACGACGACGAGAAACCTCGCGCTGCAT 903

QY 781 ATCGTTGTGGGACTCGCTGGCCATGGATGTGCTCCTCAGTCCACCAACAGACACATC 840
DB 904 ATCGTTGTGGGACTCGCTGGCCATGGATGTGCTCCTCAGTCCACCAACAGACACATC 963

QY 841 CTTGCTACTCCGCTGCTCCGCTGCGGCTACGGCTTCTACGGGGACATCATCAAGGACAGT 900
DB 964 CTTGCTACTCCGCTGCTCCGCTGCGGCTTCTACGGGGACATCATCAAGGACAGT 1023

QY 901 GAGAAGAAACGGTGTGGGTCTTGCACAGATACGACTTTTTCAGGTTTAAAGACCTTCCTC 960
DB 1024 GAGAAGAAACGGTGTGGGTCTTGCACAGATACGACTTTTTCAGGTTTAAAGACCTTCCTC 1083

QY 961 TCCACCACTGCTATGAAGGAGACAGTGTCTTCTCCTGCAACACACACCGTGGGATCT 1020
DB 1084 TCCACCACTGCTATGAAGGAGACAGTGTCTTCTCCTGCAACACACACCGTGGGATCT 1143

QY 1021 CCAAGGATAGGAGCCCTGCGGGCAGGATGCTTTGTTTTCAGGCAAGCAAGCAGCAG 1080
DB 1144 CCAAGGATAGGAGCCCTGCGGGCAGGATGCTTTGTTTTCAGGCAAGCAAGCAGCAG 1203

QY 1081 CTGAGAGGAGGACAGAAAGACACTGTAAGTGGTGGAGCTGCGGAGAGCTGAGGAG 1140
DB 1204 CTGAGAGGAGGACAGAAAGACACTGTAAGTGGTGGAGCTGCGGAGAGCTGAGGAG 1263

QY 1141 TGGCAAGTCTGTGTGGGAAAGTTTCTGGCCATCAATGCAAAACATGCTCTGTGCTTGT 1200
DB 1264 TGGCAAGTCTGTGTGGGAAAGTTTCTGGCCATCAATGCAAAACATGCTCTGTGCTTGT 1323

QY 1201 GGCGGAGCCCGGAGGGGCTCTCCCGGCTGCGGCTGCGGAGAGCGGTCTTTCAGCCTC 1260
DB 1324 GGCGGAGCCCGGAGGGGCTCTCCCGGCTGCGGCTGCGGAGAGCGGTCTTTCAGCCTC 1383

QY 1261 ATCTCTATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGACACCAAC 1320
DB 1384 ATCTCTATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGACACCAAC 1443

QY 1321 CAGCAGACCAAGTTTGAATTTTCACTTTTGTGAAGTTTATTCGCGTCAAGAAATTCAGTTT 1380
DB 1444 CAGCAGACCAAGTTTGAATTTTCACTTTTGTGAAGTTTATTCGCGTCAAGAAATTCAGTTT 1503

QY 1381 ACCTCGAAGCAGATGGAGGATGAGACAGCGACCTCAAGAGGGGGGGAAGAGCGCTTT 1440
DB 1504 ACCTCGAAGCAGATGGAGGATGAGACAGCGACCTCAAGAGGGGGGGAAGAGCGCTTT 1563

QY 1441 GGGCAGATTTTGCAGCAGCCACCTCTCTGCTGTGCAACCGTCTCCAAACAGCTCTCTGGAAC 1500
DB 1564 GGGCAGATTTTGCAGCAGCCACCTCTCTGCTGTGCAACCGTCTCTCAACAGCTCTCTGGAAC 1623
```





RESULT 8  
AB051433  
LOCUS  
DEFINITION  
AB051433  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
AB051433.1 GI:13359166  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1  
AUTHORS  
TITLE  
Hirosewa, M., Nagase, T., Murahashi, Y., Kikuno, R. and Ohara, O.  
Identification of novel transcribed sequences on human chromosome  
DNA Res. 8 (1), 1-9 (2001)  
JOURNAL  
MEDLINE  
PUBMED  
21156230  
REFERENCE  
11258795  
2 (bases 1 to 4171)  
AUTHORS  
TITLE  
Ohara, O., Nagase, T. and Kikuno, R.  
Direct Submission  
Submitted (22-NOV-2000) Osamu Ohara, Kazusa DNA Research Institute,  
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba  
292-0812, Japan (E-mail:cdmainfo@kazusa.or.jp,  
URL: http://www.kazusa.or.jp/huge, Tel: 81-438-52-3913,  
Fax: 81-438-52-3914)  
Location/Qualifiers  
1. .4171  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="hk01650"  
/note="vector: pBluescriptII SK plus"  
1. .4171  
/gene="KIAA1646"  
1. .1447  
/gene="KIAA1646"  
/note="Start codon is not identified."  
/codon\_start=2  
/product="KIAA1646 protein"  
/protein\_id="BAB33316.1"  
/db\_xref="GI:13359167"

FEATURES  
source

gene  
CDS  
ORIGIN  
Query Match 89.7%; Score 1447; DB 9; Length 4171;  
Best Local Similarity 100.0%; Pred. No. 2.4e-313;  
Matches 1447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 GATCATCGCGTGTAGGAAACAGAGCTTCACGGGAAACATCAAGCAGTGGAAATGGCA 227  
DB 1 GATCATCGCGTGTAGGAAACAGAGCTTCACGGGAAACATCAAGCAGTGGAAATGGCA 60  
QY 228 GAAATGGAAAGCCCTTACGCTTTTACAGTTCACTGTGTAAAGAGAGCAGCAGCGCACCG 287  
DB 61 GAAATGGAAAGCCCTTACGCTTTTACAGTTCACTGTGTAAAGAGAGCAGCAGCGCACCG 120  
QY 288 CTGGAAGTGGGCGCAGGTGACTTTCTGTGTCACAGAGAGCAGTGTGTCTTGTGGCT 347  
DB 121 CTGGAAGTGGGCGCAGGTGACTTTCTGTGTCACAGAGAGCAGTGTGTCTTGTGGCT 180  
QY 348 CGAGACCCCTCGGGAGATGCTGGAGAGCTGAGCTCCAGACCAAGCATTTACTGGTATT 407

DB 181 GCAGACCCCTCGGGAGATGCTGGAGAGCTGACGTCCAGACCAAGCATTTTACTGGTATT 240  
QY 408 TATCAACCCGTTTGGAGGAAAGGACAAGGCAAGCGGATATATGAAAGAAAAGTGTGCACC 467  
DB 241 TATCAACCCGTTTGGAGGAAAGGACAAGGCAAGCGGATATATGAAAGAAAAGTGTGCACC 300  
QY 468 ACTGTTACCTTAGCTCCATCACCACCTGACATCATCGTTACTTACTGAACATGCTTAATCAGC 527  
DB 301 ACTGTTACCTTAGCTCCATCACCACCTGACATCATCGTTACTTACTGAACATGCTTAATCAGC 360  
QY 528 CAAGGAGACTCTGTATGAGATTAAATAGACAAAATACGACGGCATCTGTTGTGGCGG 587  
DB 361 CAAGGAGACTCTGTATGAGATTAAATAGACAAAATACGACGGCATCTGTTGTGGCGG 420  
QY 588 AGATGGTATGTTACGAGGCTGCTGACGGTCTGATTGGGAGGACGACAGAGAGCGCCCG 647  
DB 421 AGATGGTATGTTACGAGGCTGCTGACGGTCTGATTGGGAGGACGACAGAGAGCGCCCG 480  
QY 648 GGTGACACGAGACACCCCGGCTGCTGTTGCTCCCGAGTAGCCTCCGGATTGGAATCAT 707  
DB 481 GGTGACACGAGAACCAACCCCGGCTGCTGTTGCTCCCGAGTAGCCTCCGGATTGGAATCAT 540  
QY 708 TCCCGCAGGCTCAACGAGCTGCTGTTACTTCCACCGTGGGACACAGCGACGAGAAAC 767  
DB 541 TCCCGCAGGCTCAACGAGCTGCTGTTACTTCCACCGTGGGACACAGCGACGAGAAAC 600  
QY 768 CTCGGCGTGCATATCGTTGTTGGGACCTGCTGCGCCATGAGTGTCTCAGTCCACA 827  
DB 601 CTCGGCGTGCATATCGTTGTTGGGACCTGCTGCGCCATGAGTGTCTCAGTCCACA 660  
QY 828 CAACGACACTCTCTGCTACTCGTCTCCCTGCTGGGCTACGGCTTCTACGGGACAT 887  
DB 661 CAACGACACTCTCTGCTACTCGTCTCCCTGCTGGGCTACGGCTTCTACGGGACAT 720  
QY 888 CATCAAGGACAGTGAAGAAACCGTGTGTTGGGCTCTCCAGATACGACTTTTCAAGTTT 947  
DB 721 CATCAAGGACAGTGAAGAAACCGTGTGTTGGGCTCTCCAGATACGACTTTTCAAGTTT 780  
QY 948 AAGACCTTCTCTCCACCACTGCTATGAAGGAGACAGTGTCTTCTCCCTGCAACA 1007  
DB 781 AAGACCTTCTCTCCACCACTGCTATGAAGGAGACAGTGTCTTCTCCCTGCAACA 840  
QY 1008 CACGTTGGATCTCCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTGAGCA 1067  
DB 841 CACGTTGGATCTCCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTGAGCA 900  
QY 1068 AAGCAAGCAGCAGCTGGAGGAGGACAGAAAGACACTGTATGTTTGGAGCTGCGGA 1127  
DB 901 AAGCAAGCAGCAGCTGGAGGAGGACAGAAAGACACTGTATGTTTGGAGCTGCGGA 960  
QY 1128 GGACGTGGAGGTGGCAAGTCTGTTGGGAAGTTTCTGGCCATCAATGCCACAAACAT 1187  
DB 961 GGACGTGGAGGTGGCAAGTCTGTTGGGAAGTTTCTGGCCATCAATGCCACAAACAT 1020  
QY 1188 GTCTGTGCTTGTGCGCGGAGCCCGAGGGGCTCTCCCGGCTGCCCATTTGGGAGACGG 1247  
DB 1021 GTCTGTGCTTGTGCGCGGAGCCCGAGGGGCTCTCTCCCGGCTGCCCATTTGGGAGACGG 1080  
QY 1248 GTCTTCTGACCTCATCTCCGGAATGCTCCAGGTTTCAATTTCTGAGATTTCTCAT 1307  
DB 1081 GTCTTCTGACCTCATCTCCGGAATGCTCCAGGTTTCAATTTCTGAGATTTCTCAT 1140  
QY 1308 CAGGACACCAACAGCAGGACCACTTCACTTTTGTGAAGTTTATCGGCTCAA 1367  
DB 1141 CAGGACACCAACAGCAGGACCACTTCACTTTTGTGAAGTTTATCGGCTCAA 1200  
QY 1368 GAAATTCAGTTTAGTGAAGCACTGGAGGATGAGGACGACCTCAAGAGAGGGGG 1427  
DB 1201 GAAATTCAGTTTAGTGAAGCACTGGAGGATGAGGACGACCTCAAGAGAGGGGG 1260  
QY 1428 GAAGAAGCGCTTTGGGACATTTTTCAGCAGGACCCCTCTCTGCTGCTCCAA 1487  
DB 1261 GAAGAAGCGCTTTGGGACATTTTTCAGCAGGACCCCTCTCTGCTGCTCCAA 1320

Qy 1488 CAGCTCTCTGGAATCGCAGCGGGAGGTCTCTGCACAGCCCTGCCATCGAGGTTCAGAGTCCA 1547  
Db 1371 CAGCTCTCTGGAATCGCAGCGGGAGGTCTCTGCACAGCCCTGCCATCGAGGTTCAGAGTCCA 1380  
Qy 1548 GTCCAGCTGTTTTCGACTCTTTTGCACGAGGAATGAAGAGAAATCCGAAGCCAGACTCACA 1607  
Db 1381 CTGCAGCTGGTTCGACTCTTTTGCACGAGGAATGAAGAGAAATCCGAAGCCAGACTCACA 1440  
Qy 1608 CAGCTGA 1614  
Db 1441 CAGCTGA 1447  
RESULT 9  
AX224383  
LOCUS AX224383 1840 bp DNA linear PAT 10-SEP-2001  
DEFINITION Sequence 5 from Patent WO0160990.  
ACCESSION AX224383  
VERSION AX224383.1 GI:15554633  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Rastelli, L.  
TITLE Novel sphingosine kinases  
JOURNAL Patent: WO 0160990-A 5 23-AUG-2001;  
Curagen Corporation (US); GENENTECH, INC. (US)  
FEATURES  
source 1. 1840  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 89.6%; Score 1446.4; DB 6; Length 1840;  
Best Local Similarity 97.8%; Pred. No. 4.8e-313;  
Matches 1490; Conservative 0; Mismatches 1; Indels 33; Gaps 1;  
Qy 124 GCCGCGCCCGCGCGCGATGCTGTGCTCTGTGCTGTATCTGAGATCATCGCGTTGAG 183  
Db 1 GCGCGCGCCCGCGCGCGATGCTGTGCTGTATCTGAGATCATCGCGTTGAG 60  
Qy 184 GAAACAGACGTTTCACGGGAAACATCAAGCAGTGGAAATGCGAATAATGGAAGGCT 243  
Db 61 GAAACAGACGTTTCACGGGAAACATCAAGCAGTGGAAATGCGAATAATGGAAGGCT 120  
Qy 244 TACGCTTTTACAGTTTCACTGTGTAAAGAGAGCAGCAGCGGCACCGCTGGAAGTGGCGCAG 303  
Db 121 TACGCTTTTACAGTTTCACTGTGTAAAGAGAGCAGCAGCGGCACCGCTGGAAGTGGCGCAG 180  
Qy 304 GTGACTTTTCTGGTGTCCAGAGAGCAGCTGTGTCTATTGTGGCTGACAGCCCTGCGGAG 363  
Db 181 GTGACTTTTCTGGTGTCCAGAGAGCAGCTGTGTCTATTGTGGCTGACAGCCCTGCGGAG 240  
Qy 364 ATGCTGGAGAGCTGACGCTCCAGACCAAGCAATTTACTGTTATTTATCAACCCGTTTGA 423  
Db 241 ATGCTGGAGAGCTGACGCTCCAGACCAAGCAATTTACTGGTATTTATCAACCCGTTTGA 300  
Qy 424 GGAAAGGACCAAGGCAAGCGGATATATGAAGAAAAGTGGACCACTGTTTACCTTTAGCC 483  
Db 301 GGAAAGGACCAAGGCAAGCGGATATATGAAGAAAAGTGGACCACTGTTTACCTTTAGCC 360  
Qy 484 TCCATCACACTGACATCATCG-----TTACT 510  
Db 361 TCCATCACACTGACATCATCGTAAACAAATTTCTATGTTAACTATGTAGAAATTAATTACT 420  
Qy 511 GAACATGCTTAATCAGGCCCAAGAGACTCTGTATGAGATTAAATAGACAAATACGACGC 570  
Db 421 GAACATGCTTAATCAGGCCCAAGAGACTCTGTATGAGATTAAATAGACAAATACGACGC 480

Qy 571 ATCGTCTGTGTTCGCGGAGATGTGTATTTTCAGCGAGGTGTGTGACAGGTCTGATTTGGGAGG 630  
Db 481 ATCGTCTGTGTTCGCGGAGATGTGTATTTTCAGCGAGGTGTGTGACAGGTCTGATTTGGGAGG 540  
Qy 631 ACGCAGAGGAGCGCGGGGTTCGACCAGAAACACCCCGGGCTGTGCTGGTCCCCAGTAGC 690  
Db 541 ACGCAGAGGAGCGCGGGGTTCGACCAGAAACACCCCGGGCTGTGCTGGTCCCCAGTAGC 600  
Qy 691 CTCCGGATTCGAATCATTTCCGCGAGGCTCAACGAGCTGCGTGTGTTACTCCACCGTGGGC 750  
Db 601 CTCCGGATTCGAATCATTTCCGCGAGGCTCAACGAGCTGCGTGTGTTACTCCACCGTGGGC 660  
Qy 751 ACCAGCAGCAGAAAACCTTCGCGCTGCATATCTGTTGTTGGGACTCGCTGGCCATGGAT 810  
Db 661 ACCAGCAGCAGAAAACCTTCGCGCTGCATATCTGTTGTTGGGACTCGCTGGCCATGGAT 720  
Qy 811 GTGTCTTCAGTCCACCAACAGCAGCAGCTCTTCGCTACTCCGCTGCTCCCTGCTGGGCTAC 870  
Db 721 GTGTCTTCAGTCCACCAACAGCAGCAGCTCTTCGCTACTCCGCTGCTCCCTGCTGGGCTAC 780  
Qy 871 GGCTTCTACGGGGACATCATCAAGGACAGTGAGAGAAACCGTGTGGGTCTTTGCCAGA 930  
Db 781 GGCTTCTACGGGGACATCATCAAGGACAGTGAGAGAAACCGTGTGGGTCTTTGCCAGA 840  
Qy 931 TACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAAGGACAGTGTCC 990  
Db 841 TACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAAGGACAGTGTCC 900  
Qy 991 TTCTCTCTTCGACAAACAGCAGCTGAGAGGAGCAGAGAAAGCACTGTAT 1110  
Db 901 TTCTCTCTTCGACAAACAGCAGCTGAGAGGAGCAGAGAAAGCACTGTAT 1020  
Qy 1111 GGTTTGAAGCTGCGGAGGACGTGGAGGTGGCAAGTGTCTGTGGGAAATTTCTGGCC 1170  
Db 1021 GGTTTGAAGCTGCGGAGGACGTGGAGGTGGCAAGTGTCTGTGGGAAATTTCTGGCC 1080  
Qy 1171 ATCAATGCCAACAACATGTCTGTGCTGCGCGAGCCCCAGGGGCTCTCCCCGGCT 1230  
Db 1081 ATCAATGCCAACAACATGTCTGTGCTGCGCGAGCCCCAGGGGCTCTCCCCGGCT 1140  
Qy 1231 GCCACTTTGGGAGACGGGTCTTCTGACCTCATCTCATCCGAAATGCTCCAGGTTCAAT 1290  
Db 1141 GCCACTTTGGGAGACGGGTCTTCTGACCTCATCTCATCCGAAATGCTCCAGGTTCAAT 1200  
Qy 1291 TTTCTGAGATTTCTCATCAGGCACACCAACAGCAGGACCAAGTTCACATTTTGT 1350  
Db 1201 TTTCTGAGATTTCTCATCAGGCACACCAACAGCAGGACCAAGTTCACATTTTGT 1260  
Qy 1351 GAAGTTTATCGCTCAAGAAATTCAGTTTACGTCGAGCAGCATGGAGGATGAGACAGC 1410  
Db 1261 GAAGTTTATCGCTCAAGAAATTCAGTTTACGTCGAGCAGCATGGAGGATGAGACAGC 1320  
Qy 1411 GACCTCAAGGAGGGGGGAGAAAGCGCTTTTGGGCACATTTTTCAGCAGCAGCCCTCTCTGC 1470  
Db 1321 GACCTCAAGGAGGGGGGAGAAAGCGCTTTTGGGCACATTTTTCAGCAGCAGCCCTCTCTGC 1380  
Qy 1471 TGCTGACCGTCTCCAAACAGCTCTGGAACTGCGACGGGGAGGTCTTGCAAGCCCTGCC 1530  
Db 1381 TGCTGACCGTCTCCAAACAGCTCTGGAACTGCGACGGGGAGGTCTTGCAAGCCCTGCC 1440  
Qy 1531 ATCGAGGTTCAGGTCCACTGCTGCTGCTGCTCTTTGACCGAGGAATGAGAGAAAT 1590  
Db 1441 ATCGAGGTTCAGGTCCACTGCTGCTGCTGCTCTTTGACCGAGGAATGAGAGAAAT 1500  
Qy 1591 CCGAAGCCAGACTCACACAGCTGA 1614  
Db 1501 CCGAAGCCAGACTCACACAGCTGA 1524

RESULT 10	BC067255	1772 bp	mRNA	linear	PRI 06-APR-2004
LOCUS	BC067255				
DEFINITION	Homo sapiens ceramide kinase, mRNA (cDNA clone IMAGE:6185601), complete cds.				
ACCESSION	BC067255				
VERSION	BC067255.1	GI:45595582			
KEYWORDS	.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
AUTHORS	1 (bases 1 to 1772)				
	Srausberg,R.D., Collins,F.S., Wagner,L., Shennen,C.M., Schuler,G.D., Klausner,R.D., Collins,F.S., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Altschul,S.F., Zeeberg,B., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Hopkins,R.F., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,D., Diatchenko,L., Nardusina,K., Bonaudo,M.F., Casavani,T.L., Schapleton,M., Soares,M.B., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Viallon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,M.C., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,J., Schmutz,D., Myers,R.M., Schmutz,A., Schein,J.E., Jones,S.J. and Marra,M.A.				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
REFERENCE	2 (bases 1 to 1772)				
AUTHORS	Srausberg,R.				
TITLE	Direct Submission				
JOURNAL	Submitted (12-MAR-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>				
COMMENT	Contact: MGC help desk Email: <a href="mailto:gcaps-x@mail.nih.gov">gcaps-x@mail.nih.gov</a> Tissue Procurement: Dr. James R. Lupski cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL) cDNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada <a href="mailto:info@bcgsc.bc.ca">info@bcgsc.bc.ca</a> Steve Jones, Sarah Barber, Mabel Brown-John, Varon Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu, Parvaneh Saeedi, JR Santos, Angeliq Schnerch, Ursula Skalska, Duane Smailus, Jeff Scott, Miranda Tsai, George Yang, Jacqui Schein, Asim Siddiqui, Rob Holt, Marco Marra.				
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: IRAC Plate: 141 Row: d Column: 4 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21703365 This clone has the following problem: The cds is short compared to the longest cds in the locus. Location/Qualifiers 1..1772 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6185601" /tissue_type="Peripheral Nervous System, dorsal root				
SOURCE					







Db	601	AGCAGGTCGTGTCATATGGGGTGAATTTGGGAGGACGCGAGCAGAGCGTGTGTCATCGACCCCAAT	660	AX456998.1	GI:21715789	
Qy	661	CACCCCGGGCTGTGCTGCTCCAGTAGCTCCGGAATGGAATCATTTCCCGCAGGGTCA	720			Homo sapiens (human)
Db	661	CACCCCGAGCGTGTGCTGCTCCAGTAGTACCTCAGGATCGGCATCATACCCGCGAGGGTCC	720			Homo sapiens
Qy	721	ACGGACTCGCTGTGTACTCCACCGTGGGACACGACGACGACGAGAAACCTCGCGCTGCAT	780			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Db	721	ACAGATTGTGTGTACTCAACAGTAGGCACAAACGACGACAGACATCGCTTTTGCAT	780			Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Qy	781	ATCCTTTGTTGGGACTCGCTGGCCATGAGATGTCTCAGTGCACCAACAGACACTC	840			
Db	781	ATCATATTGGGACTCATCTGCAATAGACGTCTCTGTGCACTACCAATACACGCTG	840			
Qy	841	CTTCGCTACTCCGTGCTCCTGCTGGCTACGGCTTCTACGGGGACATCATCAAGACAGT	900			
Db	841	CTGGCGTACTCGGTTCCTCTGCTGGGTACGGTTCTACGGGGACTTAATCAAGGACAGC	900			
Qy	901	GAGAAGAAACGCTGTGGTCTTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTC	960			
Db	901	GAAAGAAACGCTGTGATGGGCTCGTCGGTATGATTTCTCAGGGTTGAAGACCTTTCTC	960			
Qy	961	TCCACACACTGCTATGAAGGACAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1020			
Db	961	TCTCATCAGTACTATGAAGGACACTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1020			
Qy	1021	CCNAGGATAGGAGCCCTGCGGGCAGGATGCTTTGTTTTCAGGCAAGCAAGCAGCAG	1080			
Db	1021	CCAGGGACAATAAACCCCTGCGGGCTGGGTGCTTCTGTGTGTCAGGCGAGCAAGCAAC	1080			
Qy	1081	CTGGAGGAGGACAGAAAGACACTGTATGGTTTGAAGCTGCGGAGGAGCTCGAGGAG	1140			
Db	1081	CTGGAAGAGAGAGAGAAAGCCCTGTATGCGCTTGGAAGCGCGGAGAGTGAAGAG	1140			
Qy	1141	TGGCAAGTGTCTGTGGGAAGTTCTTGCGCCATCAATGCGCAACAATGCTCTGTCTGTGT	1200			
Db	1141	TGGCAAGTGAATGTGGGAAGTTCTTGCGCCATCAATGCGCAACAATGCTCTGTCTGTGT	1200			
Qy	1201	CGCGGAGCCCTAGGGGCTCTCCCGGCTGCCACATTTGGGAGACGGTCTCTGACCTC	1260			
Db	1201	CCTCGGAGCCCTAGGGGCTCTCCCGGCTGCCACATTTGGGAGATGGGTCTCTGACCTC	1260			
Qy	1261	ATCCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGACACCAAC	1320			
Db	1261	ATCCTTATCCGGAAGTGTCTCAGGTTCAACTTCTGAGATTTCTCATCCGSCACACGAC	1320			
Qy	1321	CAGCAGGACAGTTGACTTCACTTTTGTGAAAGTTTATCGCGTCAAGAAATTCAGTTT	1380			
Db	1321	CAGGAGGACAGTTGACTTCACTTTTGTGAAAGTTTATCGAGTCAAGAAATTCACCTT	1380			
Qy	1381	ACGTCGAGCAGATGGAAGATGAGACAAATGACTCGAAGGAACAAAGAGAGCAAGTTT	1440			
Db	1381	ACGTCGAGCAGATGGAAGATGAGACAAATGACTCGAAGGAACAAAGAGAGCAAGTTT	1440			
Qy	1441	GGGACATTTTCAGCAGCACCCCTCTCTGCTGCAACCGTCTCAACAGCTCTCTGGAAC	1500			
Db	1441	GGGAAGATCTGCAAGGACAGACCTCTTGTCACTTTGCTCAGCTTCAGNAGTCTCTGNAC	1500			
Qy	1501	TGCAGCGGGAGGTCCTGCAACGCGCTTCCATCGAGGTTCAGAGTCCACTGCGAGCTGGTT	1560			
Db	1501	TGCAGTCGGAAGTCATGCAACGCGCGCCATTGAGGTTCAGGTCCACTGCGAGCTGGTG	1560			
Qy	1561	CGACTCTTTGACGAGGAAATGAG-AGAAATCCGAGCCAGACTCACACAGC	1611			
Db	1561	CGCTCTTTTGTCTGGGGAATCGAGAAGAGTCAATGAAGCAAGACCCCAAGC	1612			
RESULT 13						
AX456998						
LOCUS						
DEFINITION						
AX456998						
ACCESSION						
Sequence 1 from Patent WO0228906.						
979 bp DNA linear PAT 06-JUL-2002						

VERSION	AX456998.1	GI:21715789				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	Kossida, S. and Encinas, J.					
AUTHORS	Regulation of human sphingosine kinase-like protein					
TITLE	Patent: WO 0228906-A 1 11-APR-2002;					
JOURNAL	Bayer Aktiengesellschaft (DE)					
FEATURES	Location/Qualifiers					
source	1..979					
	/organism="Homo sapiens"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:9606"					
ORIGIN						
Query Match	55.9%;	Score 903;	DB 6;	Length 979;		
Best Local Similarity	96.6%;	Pred. No. 1.9e-191;				
Matches	946;	Conservative 0;	Mismatches 0;	Indels 33;	Gaps 1;	
Qy	387	ACCAAAGCATTTACTGCTATTTATCAACCCGTTTGGAGGAAAGGACAAAGCCGAT	446			
Db	1	ACCAAAGCATTTACTGCTATTTTATCAACCCGTTTGGAGGAAAGGACAAAGCCGAT	60			
Qy	447	ATATGAAAGAAAGTGGCACCACCTGTTTACCTTAGCCTCCATCACCCTGACATCATCG-	505			
Db	61	ATATGAAAGAAAGTGGCACCACCTGTTTACCTTAGCCTCCATCACCCTGACATCATCG	120			
Qy	506	-----TTACTGAAACATGCTTAATCAGGCCAAAGGA	533			
Db	121	TAACAAAATTCATGTTAACTATGTAGAAGTAAATTACTGAAACATGCTTAATCAGGCCAAAGGA	180			
Qy	534	GACTCTGTATGAGATTTAATAGACAAATAGCAGCGCATGCTGTCTCGCGGAGATGG	593			
Db	181	GACTCTGTATGAGATTTAATAGACAAATAGCAGCGCATGCTGTCTCGCGGAGATGG	240			
Qy	594	TATGTTTCAGCAGAGTGTGCAACCGTCTGATTTGGAGGACGCGAGGAGCGCCGGGTCGA	653			
Db	241	TATGTTTCAGCAGAGTGTGCAACCGTCTGATTTGGAGGACGCGAGGAGCGCCGGGTCGA	300			
Qy	654	CCAGAAACACCCCGGCTGTGCTGCCAGTAGCTCCGGATTGGAATCATATCCCGC	713			
Db	301	CCAGAAACACCCCGGCTGTGCTGCCAGTAGCTCCGGATTGGAATCATATCCCGC	360			
Qy	714	AGGTCACACGACTGCTGCTTACTCCACCGTGGGACCGAGGAGCGCAGAAACCTCGGC	773			
Db	361	AGGTCACACGACTGCTGCTTACTCCACCGTGGGACCGAGGAGCGCAGAAACCTCGGC	420			
Qy	774	GCTGCATATCGTTGTTGGGACTCGCTGGGACATGGATGTGCTCCTCAGTCCACCAACACAG	833			
Db	421	GCTGCATATCGTTGTTGGGACTCGCTGGGACATGGATGTGCTCCTCAGTCCACCAACACAG	480			
Qy	834	CACACTCTTCCTGCTACTCCGTTGCTTCTGCTGGGCTACGGCTTCTACGGGGACATCAAA	893			
Db	481	CACACTCTTCCTGCTACTCCGTTGCTTCTGCTGGGCTACGGCTTCTACGGGGACATCAAA	540			
Qy	894	GGACAGTGAGAGAAACCGTGGTCTGTCAGATACGACTTTTTCAGTTTAAAGAC	953			
Db	541	GGACAGTGAGAGAAACCGTGGTCTGTCAGATACGACTTTTTCAGTTTAAAGAC	600			
Qy	954	CTTCTCTCCACACCTGCTATGAAGGGACAGTGTCTTCTCTCTGTCGACAAACACAGGT	1013			
Db	601	CTTCTCTCCACACCTGCTATGAAGGGACAGTGTCTTCTCTCTGTCGACAAACACAGGT	660			
Qy	1014	GGGATCTCCAAGGGATAGGAAGCCCTCGCGGCGAGGATGCTTTGTTTGCAGGCAAGACAA	1073			
Db	661	GGGATCTCCAAGGGATAGGAAGCCCTCGCGGCGAGGATGCTTTGTTTGCAGGCAAGACAA	720			
Qy	1074	GCAGCAGCTGGAGGAGGACGAGAGAAAGCACTGTATGTTTGGAAAGCTCGGAGGAGCT	1133			
Db	721	GCAGCAGCTGGAGGAGGACGAGAGAAAGCACTGTATGTTTGGAAAGCTCGGAGGAGCT	780			



QY 1134 GGAGAGTGGCAAGTCGTCGTGTGGGAAGTTTCTGGCCATCAATGCCACAACATGTCTCTG 1193  
DB 781 GGAGAGTGGCAAGTCGTCGTGTGGGAAGTTTCTGGCCATCAATGCCACAACATGTCTCTG 840  
QY 1194 TCGTTGTGCGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTTGGGAGACGGGTCTTC 1253  
DB 841 TCGTTGTGCGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTTGGGAGACGGGTCTTC 900  
QY 1254 TGACCTCATCTTCATCCGGAATGTCCAGGTTCAATTTCTGAGATTCTCATCAGGCA 1313  
DB 901 TGACCTCATCTTCATCCGGAATGTCCAGGTTCAATTTCTGAGATTCTCATCAGGCA 960  
QY 1314 CACCAACAGCAGGACCCAG 1332  
DB 961 CACCAACAGCAGGACCCAG 979

RESULT 14  
AK129416  
LOCUS Mus musculus mRNA for mKIAA1646 protein.  
ACCESSION AK129416  
VERSION AK129416.1 GI:37360495  
KEYWORDS FLI CDNA.  
SOURCE Mus\_musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
AUTHORS Okazaki, N., Kikuno, R., Ohara, R., Inamoto, S., Koseki, H., Hiraoka, S., Saga, Y., Nagase, T., Ohara, O. and Koga, H.  
TITLE Prediction of the coding sequences of mouse homologues of KIAA gene: III. the complete nucleotide sequences of 500 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries  
JOURNAL DNA Res. 10 (4), 167-180 (2003)  
MEDLINE 22977043  
PUBMED 14621295  
REFERENCE 2 (bases 1 to 3661)  
AUTHORS Okazaki, N., Kikuno, R., Nagase, T., Ohara, O. and Koga, H.  
TITLE Direct Submission  
JOURNAL Submitted (23-JUL-2003) Hisashi Koga, Kazusa DNA Research Institute, Laboratory for Genome Informatics; 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan (E-mail: mousekazusa.or.jp, Tel: 81-438-52-3919, Fax: 81-438-52-3918)  
COMMENT The CREATE program supported by Japan science and technology corporation; cDNA full insert sequencing: Kazusa DNA Research Institute; cDNA library construction, clone selection and 5'- & 3'-end one pass sequencing.  
FEATURES  
source Location/Qualifiers  
1..3661  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="mpn01246"  
/tissue\_type="embryonic tail"  
/note="vector:modified pBC SK+"  
1..3661  
/gene="mKIAA1646"  
41..1231  
/gene="mKIAA1646"  
/note="CDS is predicted by in silico analysis. Start codon is not identified.  
Start codon is not identified."  
/codon\_start=2  
/evidence=not\_experimental  
/product="mKIAA1646 protein"  
/protein\_id="BAC98226.1"  
/db\_xref="GI:37360496"  
translation="PCPAASRPKHLVFINPBGKGQGRIVEKTVAPLFTLASITTE  
LIITTHANQAKETLYEINTSDYGVCGGDMFSEVLHVGIGRTQQSGAIDPHNRA  
VLVPSLTLRIGIIPAGSTDCVSYVTGTDNDAETSAHLIIIGDSLVSSVHYHNTLLR

ORIGIN

Query Match 53.0%; Score 856.2; DB 10; Length 3661;  
Best Local Similarity 81.4%; Pred. No. 5.9e-181;  
Matches 1004; Conservative 0; Mismatches 228; Indels 1; Gaps 1;  
QY 380 CGTCAGACCAAGCATTTACTGTGTTATTAACAACGGTTTGGAGGAAAAGGACAAGCA 439  
DB 15 CTTCAAGACCGAAGCATTTGCTGGTATTCATCAACCTTTTCGAGGGAAGGTCAAGGCA 74  
QY 440 AGCGGATATATGAAGAAAAGTGGCACCATCTGTTACCTTAGCCTTCCATCACCCTGACA 499  
DB 75 AGCGCATCTATGAATAAACAAGTGGCGCTCTGTTTACCTTGGCTTCCATCACTACGGAGA 134  
QY 500 TCATCGTTACTGAACATGCTAATCAGGCAAGGAGACTCTGTATGAGATTAAACATAGACA 559  
DB 135 TCATCATTTACAGAGCATGCCAAACCAAGCCAAAGGAGACTTTTATACAGATCAACACAGACA 194  
QY 560 AATACGACGGCATCTGCTGCTCGCGGAGATGTTATGTTTCAGCGAGGTGCTGCACGGTC 619  
DB 195 CTTATGATGGCATCTGCTGCTAGTGGGACGGCATGTTTCAGCGAGGTGCTGCATGGGG 254  
QY 620 TGATTGGAGGACGACGAGGAGCGCGGGGTGACGACGAAACACCCCGGGCTGTGCTGG 679  
DB 255 TGATTGGAGGACGACGAGGAGCGCTGTATCGACCCCAATCACCCCGGCGCTGTGCTGG 314  
QY 680 TCCCGATGAGCTCGGAATGGAATCAATCCGAGAGGTCAACGGACTGCTGTGTTACT 739  
DB 315 TGCCCGATGAGCTCGGAATGGAATCAATCCGAGAGGTCAACGGACTGCTGTGTTACT 374  
QY 740 CCACGTTGGGACGACGAGGAGCGCGGGGTGACGACGAAACACCCCGGGCTGTGCTGG 799  
DB 375 CAACAGTGGGACGACGAGGAGCGCTGTATCGACATCATTAATTGGGGACTCGC 434  
QY 800 TGGCCATGGATGTCTCTCAGTCCACCAACAGCACATCTCTTCGCTACTCCCGTGTCCC 859  
DB 435 TGGCAATAGACGTGTCTCTGCTGCACTACCATACACGCTGCTCGGTACTCGGTTCTC 494  
QY 860 TGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGAGAGAAAACGGTGGTGG 919  
DB 495 TGCTGGGCTACGGCTTCTACGGGACATTAATCAAGGACAGCGAAAAGAACGGTGGTGG 554  
QY 920 GTCTTGGCAGATACGATTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAG 979  
DB 555 GCCTCGTCCGGTATGATTTCTCAGGGTTGAAGACCTTCTCTCATCAGTACTATGAAG 614  
QY 980 GGACAGTGTCTCTCTCCGACAAACACACGCTGGGATCTCCAAGGATAGGAAGCCCT 1039  
DB 615 GGACACTGTCTCTCTCCGACACAGCACAGCTGGGATCTCCAAGGATAGGAAGCCCT 674  
QY 1040 GCCGGGACGAGATGTTGTTTTCAGGCAAGCAAGCAGCAGCTGAGAGAGAGAGAGA 1099  
DB 675 GCCGGGCTGGGTGCTTCTGCTGACGAGCAAGCAAGCAGCTGAGAGAGAGAGA 734  
QY 1100 AAGCACTGTATGTTTGGAGCTCGGAGGAGCTGGAGGAGTGGCAAGTGGCAAGTGGGA 1159  
DB 735 AAGCCCTGTATGGCTTGGAGAACCGCGAGGAAGTGGCAAGTGGCAAGTGGGA 794  
QY 1160 AGTTTCTGGCCATCAATGGCCACAAACATGTCTCTGCTGTGCTGGCGAGGCCCGAGGGCC 1219  
DB 795 AGTTCTGGCCATCAATGGCCACCAACATGTCTCTGCTGTGCTGGCGAGCCCTGGGGCC 854  
QY 1220 TCTCCCGGCTGCCCATCTTGGGAGACGGGTCTTCTGACCTCATCTCATCGGAAATGCT 1279  
DB 855 TGTCCCATTTGCCCATCTGGGAGATGGGTCTTCTGACCTCATCTTATCCGGAAGTCT 914  
QY 1280 CCAGGTTCAATTTTCTGAGATTTCTATCAGGACACCAACACGACGACCAAGTTGACT 1339

```
Db 915 CCAGGTTCAACTCTCTGAGATTTCTCATCGGCACACGAAACAGGAGGACCAAGTTGCACT 974
Qy 1340 TCACCTTTGTTGAAGTTTATCGCTCAAGAAATTCAGTTTACCTTCGAAGCACATGAGG 1399
Db 975 TCACCTTTGTTGAAGTTTATCGAGTCAAGAAATTCACCTTCACGTCGAAGCACGTGGAAG 1034
Qy 1400 ATGAGGACAGCGACTCAAGAGGAGGGGGAAGAGCGCTTTGGGCACATTTGCAAGCAGCC 1459
Db 1035 ACAGAGCAATGACTCGAAGGAACAAGAGAGAGAGAGTTTGGGAAGATCTGCAAGGACA 1094
Qy 1460 ACCCTCTGCTGCTGCAACCGCTCTCCAAACAGCTCTCTGGAATCGCAAGGGAGGTCTGCG 1519
Db 1095 GACCCCTCTGCACTTGCTCAGCTCCAGAACTCTCTGGAATCGCAAGGGAGAGTCAATGC 1154
Qy 1520 ACAGCCCTGCACTGAGGTGAGGTCCAGCTGCGAGCTGGTTCGACTCTTTGCAAGAGGAA 1579
Db 1155 ACAGCCCGGCGCAATGAGGTGAGGTGCACTGCCAGCTGGTGGCGCTCTTTGCTCGGGAA 1214
Qy 1580 TTGAAG-AGAAATCGAAGCCAGACTCACACAGC 1611
Db 1215 TCGAGGAAGACTCATAGCAAGAACCCCAAGC 1247

RESULT 15
CR386590 1450 bp mRNA linear VRT 05-APR-2004
LOCUS Gallus gallus finished cDNA, clone ChEST291d19.
DEFINITION CR386590
ACCESSION CR386590
VERSION CR386590.1 GI:46239349
KEYWORDS
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
Boardman,P.E., Bonfield,J.K., Brown,W.R.A., Carder,C., Chalk,S.E.,
Crabing,M.D.R., Davies,R.M., Francis,M.D., Grafham,D.V.,
Hubbard,S.J., Humphray,S.J., Hunt,P.J., Maddison,M., McLaren,S.R.,
Niblett,D., Overton,I.M., Rogers,J., Scott,C.E., Taylor,R.G.,
Tickle,C. and Wilson,S.A.
Direct Submission
Submitted (05-APR-2004) Sanger Institute, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: chickest@bms.umist.ac.uk
BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA
sequencing project.
This sequence is from the
BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,
from a library constructed by Elizabeth Bosch. cDNA was prepared
from RNA extracted from heads, normalised, and poly A-trimmed.
EcoRI-NotI cut cDNA was then ligated into the vector. Vector:
pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host: Escherichia
coli DH10B.
FEATURES
Location/Qualifiers
source 1..1450
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, HiseX"
/db_xref="taxon:9031"
/clone="ChEST291d19"
/clone_lib="CSEQCHN23"
/dev_stage="stage 22"

ORIGIN
Query Match 40.7%; Score 656.8; DB 5; Length 1450;
Best Local Similarity 68.2%; Pred. NO. 2.6e-136;
Matches 946; Conservative 0; Mismatches 432; Indels 10; Gaps 2;

Qy 202 AAACATCAAGCGAGTGGAAATGGCAGAAATGGAAAGCCCTTACGCTTTTACAGTTTCAAC 261
Db 1 AAACAGCGTAGTATTTGGCAATGGCAAAAGATGGCAAGCCCAAGTCTTTCACAGTTTAT 60
Qy 262 TGTGTAAAGAGACGACGCGGACCGCTGGAGTGGGCGGAGGTGACTTTTCTGGTGTCCA 321
```

```
Db 61 TACGTGAAAAAGCCCGAAATCACCGCTGGCGGTGCAGAGATGTGACGTTTGGTGTGCT 120
Qy 322 GAGGAGCAGCTGTGTCACTTGTGCTGCACACCTCGCGGAGAGATGCTGGAGAGCTCAGC 381
Db 121 GATGAGATCTTGTGTAAACCACTGGATACAGGCACT----GAATTACTTTGAAATGCAGAC 176
Qy 382 TCAGAGACAAAGCATTTACTTGTGATTTATCAACCCGTTTGGAGGAAAGCAAGGCAAG 441
Db 177 TGTAGACCAAGCAGTTGCTTGTGATATTAATCCATATATGAGGAAACCAAGAGGAG 236
Qy 442 CGGATATATGAAGAAAGTGGCAACACTGTTCCACTTAGCTTCCATCACACTGACATC 501
Db 237 AGGATTTATGAACAGAAAGTTGCTCCACTCTTCAGCTTGGCTTCTATTCTTACTGATGTT 296
Qy 502 ATCGTTACTCAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAA 561
Db 297 GTCGTAACATGAATGCTAACCACTGCTAAGGACAACCTTATTTGAAGTTAATATTAATAA 356
Qy 562 TAGCAGCGCATCGTCTGTGTCGCGGAGATGCTGATGTTTCAGCGAGGTGCTGCACGCTG 621
Db 357 TATGATGCTGTGTTGTTGTTGTTGGGACGGCATGTTCACTGAAGTATGATGCTGCTC 416
Qy 622 ATTGGAGAGACGAGAGAGCGCGCGGGGTGCAACAGAAACCAACCCCGGGCTGTGCTGGTC 681
Db 417 ATTGGAAGAAATGCAGAGAGGACTCTGGCATAGACCAAAATAATCCCAAAGCACCGTTAGTC 476
Qy 682 CCCAGTAGCCTCCGGATTGGAATCATTTCCCGCAGGGTCAACGGAATCGCTGTGTTACTTCC 741
Db 477 CAGTCAATATTAAGGATTGGCATAAATTCCTGCTGGCTCAACAGATTCCGATGATGTAITCA 536
Qy 742 ACCGTGGGACACGAGCGACGAGAAACCTCGCGCTGCATATCTGTTGTTGGGACTCGCTG 801
Db 537 ACTGTTGGCATTTCTGATCCAGTAACATCAGCTCTTCATATTAATATAGGTGACTGTGAG 596
Qy 802 GCCATGATGTGCTCTCAGTCCACCAACAAGCACTCCTTCCTTCTGCTACTCGTGTCCCTG 861
Db 597 CTTTATGATGCTCATCTGTGTCATCAAAACAACACGTTTGTGAAGTATGCTGTATCATTTG 656
Qy 862 CTGGGCTACGGCTTCTACGGGGACATCATCAAGGACAGTGGAGAGAAACGTTGTTGGGT 921
Db 657 TTGGGCTATGGTTTTATGAGAGATGTTTTGAAGACAGTGAAGAAACGTTGATGGGT 716
Qy 922 CTTGCCAGATACGACTTTTCAGAGTTTAAAGACCTTCCTCTCCCACTGCTGTATGAAGGG 981
Db 717 CCGATGAGATACGACTATTCAAGCTTCAAGACTTTTCTTCTCATCTACTATGAAGGA 776
Qy 982 ACAGTGTCTTCTCTCTGCAACAACAACGCGTGGGATCTCCAAAGGATAGGAAGCCCTGC 1041
Db 777 ACAATTTCTTTTCAACAGCAAAACACACACTGGGATCTCCACGAGATAAAGATAGCTGC 836
Qy 1042 CGGGCAGGATGCTTTGTTTCAGCGCAAGCAAGCAGCAGCTGGAGGAGGAGCAGCAAGAAA 1101
Db 837 AGAACAGGATGTTATATTTGCAAGAAAGTGAGCAACAACCTGGCAGAACACGCAAGAA- 895
Qy 1102 GCACTGTATGTTTGGAGAGTTCGCGAGGAGCTGGGAGGTGGCAAGTCTGCTGTGGGAAG 1161
Db 896 -----GTGTGGATTCAACATGAAGAGATGAAGAAGTGAAGGTTATTAAAGGGAAA 950
Qy 1162 TTTCTGCGCATCAATGCGCAAAACATGCTCTGTGCTGTTGTGCGCGAGGCCCGAGGGGCTC 1221
Db 951 TTTCTAGCCATCAATGTAGTAAATATGTGTGTGCTGTCTGCCACGAAGTCCAAAAAGGTCTT 1010
Qy 1222 TCCCGCGCTGCCACTTTGGGAGACGGGTCTTCTGACCTCATCTCTCATTCGGAATGTCTCC 1281
Db 1011 TCACCGGAGCTCAATTTAGCAGATGGTTAGCTGACCTGATCTTAGTTCGGAATGTCTCC 1070
Qy 1282 AGGTTCAATTTCTGAGATTTCTCATCAGGCACACCAACCAAGCAGGAGCCAGTTTACTTTC 1341
Db 1071 AGATTTGATTTCTGCGTTATCTTGTGACATACAAACAAGATGACCAAGTTTACTTTC 1130
Qy 1342 ACTTTTGTGAAGTTTATCGGTCAGAAATTCAGTTTACGTTAGCTGGAAGCATCGGAGT 1401
```

Db 1131 TCATTTGTAGATGTTTATCGGCTGAAGAGTTTTCAATTTACATCAAAGCTTTCAGAAGAC 1190  
QY 1402 GAGGACAGCGACCTCAAGAGGGGGGGAAGAGCGCTTTGGGCACATTTGCAGCAGCCAC 1461  
Db 1191 AATGAAAGCAGTGTACACAGACATAGGAAAGAAACATTTTGGCCAGTTCTGCAGAGATCAT 1250  
QY 1462 CCTCCTGTGTGTGCACCGTCTCCAAACAGCTCCTGGAACGTGCGCGGGAGGTCTCTGCAC 1521  
Db 1251 CCAGCTGTCTGTCAATATTGCTAATAGCACTTGGAAATTGTATGGAGAAACCTTGGAT 1310  
QY 1522 AGCCCTGCCATCGAGGTCAAGTCCACTGCCAGCTGGTTCCGACTCTTTGCACGAGGAATT 1581  
Db 1311 AGCTCAGCAATTGAAGTGAGGGTTTCACTGCCAGTTAATGAAACTGTTTTGCAGGGGAATC 1370  
QY 1582 GAAGAGAA 1589  
Db 1371 GAGGAAAA 1378

Search completed: September 5, 2005, 23:41:46  
Job time : 7343.65 secs

**This Page Blank (uspto)**





KW lipid kinase; KIAA1646.  
 XX  
 OS Homo sapiens.  
 OS 39.  
 XX FH Key Location/Qualifiers  
 XX FT variation replace(2391,g)  
 XX FT /tag= a  
 XX FT /standard\_name= "Single nucleotide polymorphism"  
 XX FT replace(2577,g)  
 XX FT /tag= b  
 XX FT /standard\_name= "Single nucleotide polymorphism"  
 XX FT variation replace(3020,c)  
 XX FT /tag= c  
 XX FT /standard\_name= "Single nucleotide polymorphism"  
 XX FT variation replace(3769,c)  
 XX FT /tag= d  
 XX FT /standard\_name= "Single nucleotide polymorphism"  
 XX FT variation replace(4272,g)  
 XX FT /tag= e  
 XX FT /standard\_name= "Single nucleotide polymorphism"  
 XX  
 PN WO2004006838-A2.  
 XX  
 XX  
 XX  
 PD 22-JAN-2004.  
 XX  
 PF 15-JUL-2003; 2003WO-US021730.  
 XX  
 XX  
 PR 15-JUL-2002; 2002US-0395632P.  
 XX  
 XX (SUGEN-) SUGEN INC.  
 XX  
 XX Whyte D, Manning G, Caenepeel S;  
 PI WPI; 2004-122753/12.  
 DR P-PSDB; ADJ96664.  
 XX  
 PT New nucleic acid molecule encoding a kinase polypeptide, useful for  
 PT preparing a composition for treating diseases or disorders, e.g., cancer,  
 PT or neurological, immunological or inflammatory disorders.  
 XX  
 XX Example 1; SEQ ID NO 55; 366pp; English.  
 PS  
 XX This invention relates to a novel isolated, enriched or purified nucleic  
 CC acid molecule that encodes a kinase polypeptide. Specifically, it relates  
 CC to human tyrosine and serine/threonine protein kinases (PK's and STK's),  
 CC as well as protein kinase-like enzymes. The present invention describes  
 CC screening methods to identify agonists, antagonists and antibodies that  
 CC can be used to modulate the activity or function of the mammalian kinase  
 CC enzymes. As such, these compositions can be used for gene therapy  
 CC purposes to treat diseases or disorders including cancer, immune-related  
 CC diseases, cardiovascular disease, brain or neuronal associated disease,  
 CC metabolic and inflammatory disorders. Accordingly, they exhibit  
 CC cytostatic, neuroprotective, immunomodulator and antiinflammatory  
 CC activities. This polynucleotide sequence is a human kinase DNA sequence  
 CC of the invention.  
 XX  
 SQ Sequence 4429 BP; 1020 A; 1115 C; 1213 G; 1081 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 1614; DB 12; Length 4429;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1614; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGGGGGCGACGGGGCGGCGGCGGCTGCAATCCGTGCTGGGTGAAGCAGCAGCGC 60  
 DB 92 ATGGGGCGACGGGGCGGCGGCGGCTGCAATCCGTGCTGGGTGAAGCAGCAGCGC 151  
 QY 61 TCGCCGCTGAGCTGAGAGCCCGCGGCTCTGCTGCGCTGCTGCGGAGCCCGGGGCC 120  
 DB 152 TCGCCGCTGAGCTGAGAGCCCGCGGCTCTGCTGCGCTGCTGCGGAGCCCGGGGCC 211  
 QY 121 GGAGCCGGCCCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
 |||

Db 212 GGAGCGGCGCCCGCGCGGATGCCCTGCTGTGCTGTATCTGAGATCATCGCCGTT 271  
 QY 181 GAGGAAA CAGACGTTTACGGGAAAACATCAAGGACGTGGAATATGCAAGAAAAG 240  
 DB 272 GAGGAAA CAGACGTTTACGGGAAAACATCAAGGACGTGGAATATGCAAGAAAAG 331  
 QY 241 CCTTACGCTTTTACAGTTTCACTGTGTAAGAGAGACGACGCGCAGCTGGAAGTGGCG 300  
 DB 332 CCTTACGCTTTTACAGTTTCACTGTGTAAGAGAGACGACGCGCAGCTGGAAGTGGCG 391  
 QY 301 CAGGTGACTTTCTGGTGTCCAGAGGAGCAGCTGTGTCACTTGTGTGCTGCAGACCTTGC 360  
 DB 392 CAGGTGACTTTCTGGTGTCCAGAGGAGCAGCTGTGTCACTTGTGTGCTGCAGACCTTGC 451  
 QY 361 GAGATGCTGGAGAAAGCTGACGTCCAGACCAAGCAATTTTCTGTGTTATTTATCAACCGTTT 420  
 DB 452 GAGATGCTGGAGAAAGCTGACGTCCAGACCAAGCAATTTTCTGTGTTATTTATCAACCGTTT 511  
 QY 421 GAGGAAAAGGACCAAGGCAAGGGATATATGAAGAAAAGTGGACCACTGTTTCACTT 480  
 DB 512 GAGGAAAAGGACCAAGGCAAGGGATATATGAAGAAAAGTGGACCACTGTTTCACTT 571  
 QY 481 GCCTCCATCACCACCTGACATCATCGTTACTGAAATGCTTAATCAGGCCAAGGAGACTCTG 540  
 DB 572 GCCTCCATCACCACCTGACATCATCGTTACTGAAATGCTTAATCAGGCCAAGGAGACTCTG 631  
 QY 541 TATGAGATTAAATAGACAAATAACGCGCATTCGCTGTGTGCGGCGAGATGGTATGTTTC 600  
 DB 632 TATGAGATTAAATAGACAAATAACGCGCATTCGCTGTGTGCGGCGAGATGGTATGTTTC 691  
 QY 601 AGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACGAGAGCGCGGGTGCACGAGAAC 660  
 DB 692 AGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACGAGAGCGCGGGTGCACGAGAAC 751  
 QY 661 CACCCCGGGCTGTGCTGCTCCAGTAGCTCCGGAATTGGAATCATTTCCCGCAGGCTCA 720  
 DB 752 CACCCCGGGCTGTGCTGCTCCAGTAGCTCCGGAATTGGAATCATTTCCCGCAGGCTCA 811  
 QY 721 ACGAGTCTGCTGTGTTTACTCCACGTTGGGACACGACGACGAGAACTCTGCGCTGCAAT 780  
 DB 812 ACGGACTGCGTGTGTTTACTCCACGTTGGGACACGACGACGAGAACTCTGCGCTGCAAT 871  
 QY 781 ATCGTTGTTGGGACTCGCTGCCCATGATGTGCTCTCAGTCCACCAACAGCACACTC 840  
 DB 872 ATCGTTGTTGGGACTCGCTGCCCATGATGTGCTCTCAGTCCACCAACAGCACACTC 931  
 QY 841 CTTGCTACTCCGTGCTCCCTGCTGGCTTACGGGATCATCATCAAGCAGACT 900  
 DB 932 CTTGCTACTCCGTGCTCCCTGCTGGCTTACGGGATCATCATCAAGCAGACT 991  
 QY 901 GAGAAGAAACGGTGGTGGGCTTTGCCAGATACGATTTTTCAGGTTTAAAGACCTTCTCTC 960  
 DB 992 GAGAAGAAACGGTGGTGGGCTTTGCCAGATACGATTTTTCAGGTTTAAAGACCTTCTCTC 1051  
 QY 961 TCCACCACTCTATGAAGGACAGTGTCTTCTCCTCTGCACACACACGTTGGGATCT 1020  
 DB 1052 TCCACCACTCTATGAAGGACAGTGTCTTCTCCTCTGCACACACACGTTGGGATCT 1111  
 QY 1021 CCAAGGATAGAAAGCCCTCGCGGACAGTGTGTTTTCAGGCAAGCAAGCAGCAG 1080  
 DB 1112 CCAAGGATAGAAAGCCCTCGCGGACAGTGTGTTTTCAGGCAAGCAAGCAGCAG 1171  
 QY 1081 CTGGAGGAGGACAGAAAGACACTGTATGTTTGAAGCTTGGAGGAGAGCTGGAGGAG 1140  
 DB 1172 CTGGAGGAGGACAGAAAGACACTGTATGTTTGAAGCTTGGAGGAGAGCTGGAGGAG 1231  
 QY 1141 TGGCAAGTCTGTGGGAAGTTTCTGGCCATCAATGCCAACAACATGCTCTGTGCTTGT 1200  
 DB 1232 TGGCAAGTCTGTGGGAAGTTTCTGGCCATCAATGCCAACAACATGCTCTGTGCTTGT 1291  
 QY 1201 CGCCGAGCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGAGCGGCTCTTCTGACCTC 1260  
 DB 1292 CGCCGAGCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGAGCGGCTCTTCTGACCTC 1351





1804	Db	AGCAGAGTGTGCACGGTCTCTGATTGGGAGAGACGCAGAGAGGCGCGCGGGTTCGACCAAGAAC	1863
661	Qy	CACCCCGGGGTGTGCTGGTCCCGCAGTAGCCTCCGGATTGGAAATCAATTCGCGCAGGGTCA	720
1864	Db	CACCCCGGGGTGTGCTGGTCCCGCAGTAGCCTCCGGATTGGAATCAATTCGCGCAGGGTCA	1923
721	Qy	ACGACACTGCGTGTGTTACTCACCGTGGGACACAGGACGCAGAGAAACCTTGGCGCGTGCAT	780
1924	Db	ACGACACTGCGTGTGTTACTCACCGTGGGACACAGGACGCAGAGAAACCTTGGCGCGTGCAT	1983
781	Qy	ATCCGTGTTGGGGACTCGCTGGCCATGGATGTGTCTCAGTCCACCAACAACAGCACACTC	840
1984	Db	ATCCGTGTTGGGGACTCGCTGGCCATGGATGTGTCTCAGTCCACCAACAACAGCACACTC	2043
841	Qy	CTTCGCTACTCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGT	900
2044	Db	CTTCGCTACTCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGT	2103
901	Qy	GAGAAAGAACGGTGGTGGGTCTTCCAGATACACACTTTTCAGGTTTAAAGACCTTCCTC	960
2104	Db	GAGAAAGAACGGTGGTGGGTCTTCCAGATACACACTTTTCAGGTTTAAAGACCTTCCTC	2163
961	Qy	TCCCAACCACTGCTATGAAGGACAGTGTCTCTCTCCTGCAACAACACAGTGGGATCT	1020
2164	Db	TCCCAACCACTGCTATGAAGGACAGTGTCTCTCTCCTGCAACAACACAGTGGGATCT	2223
1021	Qy	CCAAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTGCAGGCAAAAGCAAGCAGCAG	1080
2224	Db	CCAAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTGCAGGCAAAAGCAAGCAGCAG	2283
1081	Qy	CTGGAGGAGGACAGAAAGACACTGTATGGTTTTGAAGCTGCGGAGGAGCTGGAGGAG	1140
2284	Db	CTGGAGGAGGACAGAAAGACACTGTATGGTTTTGAAGCTGCGGAGGAGCTGGAGGAG	2343
1141	Qy	TGGCAAGTCTGTCTGGGAAGTCTTCGGCCATCAATGCCACAACAATGTCTCTGTGCTTGT	1200
2344	Db	TGGCAAGTCTGTCTGGGAAGTCTTCGGCCATCAATGCCACAACAATGTCTCTGTGCTTGT	2403
1201	Qy	CGCGGAGCCCCCAGGGCCCTCTCCCCGGCTGCCCACTTGGGAGACGGGTCTTCTGACCTC	1260
2404	Db	CGCGGAGCCCCCAGGGCCCTCTCCCCGGCTGCCCACTTGGGAGACGGGTCTTCTGACCTC	2463
1261	Qy	ATCCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAAC	1320
2464	Db	ATCCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAAC	2523
1321	Qy	CAGCAGACCACTTTGACTTCACTTTTGTGTAAGTTTATCCGCTCAAGAAATTCACGTTT	1380
2524	Db	CAGCAGACCACTTTGACTTCACTTTTGTGTAAGTTTATCCGCTCAAGAAATTCACGTTT	2583
1381	Qy	ACGTGCAAGCACATGGAGGATGAGGACAGCGACCTCAAGGAGGGGGGGAAGAACGCTTT	1440
2584	Db	ACGTGCAAGCACATGGAGGATGAGGACAGCGACCTCAAGGAGGGGGGGAAGAACGCTTT	2643
1441	Qy	GGGCACATTTGACAGCCACCCCTCTGCTGTGCACCGTCTCCACAGCTCCTGTGAAC	1500
2644	Db	GGGCACATTTGACAGCCACCCCTCTGCTGTGCACCGTCTCCACAGCTCCTGTGAAC	2703
1501	Qy	TGCGACGGGAGGTCTTGCAACAGCCCTGCCATCGAGGTTCAGAGTCCACTGCGCAGCTGTT	1560
2704	Db	TGCGACGGGAGGTCTTGCAACAGCCCTGCCATCGAGGTTCAGAGTCCACTGCGCAGCTGTT	2763
1561	Qy	CGACTCTTTTGCACGAGGAATTGAAGAAATCCGAAGCCACACTCACAACGCTGA	1614
2764	Db	CGACTCTTTTGCACGAGGAATTGAAGAAATCCGAAGCCACACTCACAACGCTGA	2817

RESULT 4  
ADP55247  
ID ADP  
XX  
AC ADP  
XX

DT	18-NOV-2004	(first entry)
DE	Human PRO cDNA sequence SEQ ID NO:1223.	
KW	human; PRO; immune related disease; inflammatory immune response;	
KW	immune response stimulation; anti-allergic; antianaemic; antiarthritic;	
KW	antiasthmatic; antidiabetic; antinflammatory; antipariatic;	
KW	antirheumatic; antithyroid; CNS; dermatological; gastrointestinal;	
KW	haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;	
KW	nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;	
KW	virucide; gene therapy; gene; ss.	
OS	Homo sapiens.	
PN	WO2004039956-A2.	
XX		
XX	13-MAY-2004.	
PD		
XX		
PF	28-OCT-2003; 2003WO-US034381.	
XX		
XX	29-OCT-2002; 2002US-0422472P.	
PR		
XX		
PA	(GETH ) GENENTECH INC.	
XX		
XX	Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;	
PI	Wood WI, Wu TD;	
FI		
XX		
XX	WPI; 2004-376182/35.	
DR		
DR	P-PSDB; ADF55248.	
XX		
PT	New PRO polynucleotides and polypeptides, useful in diagnosing	
PT	and treating an immune related disease, e.g. systemic lupus	
PT	erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in	
PT	stimulating an immune response.	
XX		
XX	Claim 2; SEQ ID NO 1223; 3009pp; English.	
PS		
XX		
CC	The present invention describes an isolated PRO nucleic acid (1). Also	
CC	described: (1) a vector comprising (1); (2) a host cell comprising the	
CC	vector of (1); (3) a process for producing a PRO polypeptide; (4) an	
CC	isolated PRO polypeptide; (5) a chimeric molecule comprising the	
CC	polypeptide of (4) fused to a heterologous amino acid sequence; (6) an	
CC	antibody which specifically binds to a polypeptide of (4); (7) a	
CC	composition of matter comprising a polypeptide of (4), an agonist or	
CC	antagonist of the polypeptide or an antibody that binds to the	
CC	polypeptide in combination with a carrier; (8) an article of manufacture	
CC	comprising a container, a label on the container and a composition of	
CC	matter of (7); (9) a method of treating an immune related disease in a	
CC	mammal; (10) a method for determining the presence of a PRO polypeptide	
CC	in a sample suspected of having the polypeptide; (11) a method of	
CC	diagnosing an immune related disease or an inflammatory immune response	
CC	in mammal; (12) a method of identifying a compound that inhibits or	
CC	inactivates the activity of or expression of a gene encoding a PRO polypeptide	
CC	; and (13) a method of stimulating the immune response in a mammal. The	
CC	PRO sequences have anti-allergic, antianaemic, antiarthritic,	
CC	antiasthmatic, antidiabetic, antinflammatory, antipariatic,	
CC	antirheumatic, antithyroid, CNS, dermatological, gastrointestinal,	
CC	haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,	
CC	nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and	
CC	virucide activities, and can be used in gene therapy. The nucleic acid	
CC	(1) and the encoded polypeptides, compositions, kits and methods are	
CC	useful in diagnosing and treating an immune related disease and in	
CC	stimulating an immune response. The present sequence represents a human	
CC	PRO nucleotide sequence from the present invention.	
XX		
XX	Sequence 4445 BP; 1018 A; 1134 C; 1216 G; 1077 T; 0 U; 0 Other;	

```

Query Match      99.9%; Score 1612.4; DB 13; Length 4445;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1613; Conservative 0; Mismatches 1; Indels 0; Gaps 0

Qy      1  ATGGGGGCGACGGGGGGGGGACCGCGTGCATCCGTGCTGGGTGAAGCAGCAGCGC 60

```

Db 124 ATGGGGCGACGGGGCGCGAGCCGCTGCAATCCGTGCTGTGGGTGAAGCAGCAGCGC 183  
Qy 61 TGGCGCGTGAAGCTGGAGCCCGCGGCGCTCTGCTGGCGTGTGGCGAGCCCGGGGCC 120  
Db 184 TGGCGCGTGAAGCTGGAGCCCGCGGCGCTCTGCTGGCGTGTGGCGAGCCCGGGGCC 243  
Qy 121 GGAGCGCGCGCCCGCGGCGGAGTCCCTGCTCTGTGCTGTATCTGAGATCATCGCGTT 180  
Db 244 GGAGCGCGCGCCCGCGGCGGAGTCCCTGCTCTGTGCTGTATCTGAGATCATCGCGTT 303  
Qy 181 GAGGAACAGACGTTTCAGGGAAACATCAAGGCGAGTGAAGAAATGGCAGAAATGGAAG 240  
Db 304 GAGGAACAGACGTTTCAGGGAAACATCAAGGCGAGTGAAGAAATGGCAGAAATGGAAG 363  
Qy 241 CCTTACGCTTTTACAGTTTCACTGTGTAAGAGAGCAGCAGCGCACCGCTGGAAGTGGCG 300  
Db 364 CCTTACGCTTTTACAGTTTCACTGTGTAAGAGAGCAGCAGCGCACCGCTGGAAGTGGCG 423  
Qy 301 CAGGTGACTTTTCTGGTGTCCAGAGAGCAGCTGTGTCACTTGTGGCTGCAGACCTGCGG 360  
Db 424 CAGGTGACTTTTCTGGTGTCCAGAGAGCAGCTGTGTCACTTGTGGCTGCAGACCTGCGG 483  
Qy 361 GAGATGCTGGAGAGCTGACGTCAGACCAAGCATTTTACTGGTATTTATCAACCCGTTT 420  
Db 484 GAGATGCTGGAGAGCTGACGTCAGACCAAGCATTTTACTGGTATTTATCAACCCGTTT 543  
Qy 421 GGAGAAAGACAGGCAAGCGGATATATGAAGAAAGTGGCACCACCTGTCACCTTA 480  
Db 544 GGAGAAAGACAGGCAAGCGGATATATGAAGAAAGTGGCACCACCTGTCACCTTA 603  
Qy 481 GCCTCCATCACCACCTGACATCATCGTTACTGAAACATGCTAATCAGGCGCAAGGAGACTGTG 540  
Db 604 GCCTCCATCACCACCTGACATCATCGTTACTGAAACATGCTAATCAGGCGCAAGGAGACTGTG 663  
Qy 541 TATGAGATTAAATAGACAAATACAGCGCATCGTCTGTGTGGCGGAGATGTTATGTTTC 600  
Db 664 TATGAGATTAAATAGACAAATACAGCGCATCGTCTGTGTGGCGGAGATGTTATGTTTC 723  
Qy 601 AGCGAGGTGCTGACGGTCTGATTGGGAGGACGACAGAGCGCGCGGGTGCACAGAAC 660  
Db 724 AGCGAGGTGCTGACGGTCTGATTGGGAGGACGACAGAGCGCGCGGGTGCACAGAAC 783  
Qy 661 CACCCCGGGTGTGCTGGTCCCCAGTAGCTCCGGATTGGAATCATTTCCCGCAGGGTCA 720  
Db 784 CACCCCGGGTGTGCTGGTCCCCAGTAGCTCCGGATTGGAATCATTTCCCGCAGGGTCA 843  
Qy 721 ACGGACTCGGTGTGTTACTCCACCGTGGGACACGACGACGAGAACTCCGCGCTGCAT 780  
Db 844 ACGGACTCGGTGTGTTACTCCACCGTGGGACACGACGACGAGAACTCCGCGCTGCAT 903  
Qy 781 ATCGTTGTTGGGACTCGCTGGCCATGGATGTGTCCTCAGTCCACCAACAGACACATC 840  
Db 904 ATCGTTGTTGGGACTCGCTGGCCATGGATGTGTCCTCAGTCCACCAACAGACACATC 963  
Qy 841 CTTGCTACTCCGTGCTCCTGCTGGGCTACCGGCTTCTACGGGGACATCATCAAGACAGT 900  
Db 964 CTTGCTACTCCGTGCTCCTGCTGGGCTACCGGCTTCTACGGGGACATCATCAAGACAGT 1023  
Qy 901 GAGAAAGAACGGTGTGGGCTTGGCCAGATACGACTTTTTCAGTTTAAAGACCTTCCTC 960  
Db 1024 GAGAAAGAACGGTGTGGGCTTGGCCAGATACGACTTTTTCAGTTTAAAGACCTTCCTC 1083  
Qy 961 TCCACCACTGCTATGAAGGACAGTGTCTTCTCCCTGCACACACACACCGTGGGATCT 1020  
Db 1084 TCCACCACTGCTATGAAGGACAGTGTCTTCTCCCTGCACACACACACCGTGGGATCT 1143  
Qy 1021 CCAAGGATAGGAAGCCCTGCGGGCAGAGATGCTTTGTTTGCAGGCAAGCAAGCAGCAG 1080  
Db 1144 CCAAGGATAGGAAGCCCTGCGGGCAGAGATGCTTTGTTTGCAGGCAAGCAAGCAGCAG 1203  
Qy 1081 CTGGAGGAGGACAGCAAGAACACTGTATGGTTTGGAAAGTGGCGGAGAGTGGAGAG 1140  
Db 1204 CTGGAGGAGGACAGCAAGAACACTGTATGGTTTGGAAAGTGGCGGAGAGTGGAGAG 1263

Qy 1141 TGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCCAAAACATGTCTGTGCTTGT 1200  
Db 1264 TGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCCAAAACATGTCTGTGCTTGT 1323  
Qy 1201 CGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACATTGGGAGACGGGTCTTCTGACCTC 1260  
Db 1324 CGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACATTGGGAGACGGGTCTTCTGACCTC 1383  
Qy 1261 ATCCTCATCCGGAATGCTCCAGGTTTCAATTTTCTGAGATTTCTCATCAGGCACCAAC 1320  
Db 1384 ATCCTCATCCGGAATGCTCCAGGTTTCAATTTTCTGAGATTTCTCATCAGGCACCAAC 1443  
Qy 1321 CAGCAGGACCAAGTTTGAATTTTCACTTTTGTGAAAGTTTATCGCGTCAAGAAATCCAGTTT 1380  
Db 1444 CAGCAGGACCAAGTTTGAATTTTCACTTTTGTGAAAGTTTATCGCGTCAAGAAATCCAGTTT 1503  
Qy 1381 ACCTCGAAGCACATGGAGGATGAGGACAGCGACTCAAGAGGGGGGGAAGACGCTTT 1440  
Db 1504 ACCTCGAAGCACATGGAGGATGAGGACAGCGACTCAAGAGGGGGGGAAGACGCTTT 1563  
Qy 1441 GGGCACATTTGCAGCAGCCCTCTGCTGTGTCACCGTCTCCAAACAGCTCTGGAAC 1500  
Db 1564 GGGCACATTTGCAGCAGCCCTCTGCTGTGTCACCGTCTCCAAACAGCTCTGGAAC 1623  
Qy 1501 TGGCAGCGGAGGTCTCTGCACAGCCCTGCCATCGAGGTCAGAGTCCACTGCCAGCTGGTT 1560  
Db 1624 TGGCAGCGGAGGTCTCTGCACAGCCCTGCCATCGAGGTCAGAGTCCACTGCCAGCTGGTT 1683  
Qy 1561 CGACTCTTTGACAGAGAAATTGAAGAGAAATCCGAAGCCAGACTCACACAGCTGA 1614  
Db 1684 CGACTCTTTGACAGAGAAATTGAAGAGAAATCCGAAGCCAGACTCACACAGCTGA 1737  
  
RESULT 5  
ADA05679  
ID ADA05679 standard; cDNA; 1740 BP.  
XX  
AC ADA05679;  
XX  
XX  
DT 06-NOV-2003 (first entry)  
XX  
DE Human NOV9a encoding cDNA SEQ ID NO:39.  
XX  
KW human; NOV9; antidiabetic; anorectic; antibacterial; virucide;  
KW immunomodulator; cytosolic; nootropic; neuroprotective;  
KW antiparkinsonian; antilipemic; gene therapy; human disease;  
KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW immune disorder; haematopoietic disorder; dyslipidaemia; gene; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 76..1689  
FT /\*tag= a  
FT /product= "NOV9a"  
XX  
XX WO2003029424-A2.  
XX  
PD 10-APR-2003.  
XX  
PF 02-OCT-2002; 2002WO-US031373.  
XX  
PR 02-OCT-2001; 2001US-0326483P.  
PR 05-OCT-2001; 2001US-0327435P.  
PR 05-OCT-2001; 2001US-0327449P.  
PR 09-OCT-2001; 2001US-0327917P.  
PR 09-OCT-2001; 2001US-0328029P.  
PR 09-OCT-2001; 2001US-0328044P.  
PR 09-OCT-2001; 2001US-0328056P.  
PR 12-OCT-2001; 2001US-0328849P.  
PR 15-OCT-2001; 2001US-0329414P.



```
Qy 961 TCCACCACTGCTATGAGGACAGTGTCTTCTCTCCCTGCAACAACACAGGTGGATCT 1020
Db 1036 TCCACCACTGCTATGAGGACAGTGTCTTCTCTCCCTGCAACAACACAGGTGGATCT 1095
Qy 1021 CCAAGGATAGAACCCCTGCCGGCAGGATGCTTTGTTTCAGGCAAGCAAGCAGCAG 1080
Db 1096 CCAAGGATAGAACCCCTGCCGGCAGGATGCTTTGTTTCAGGCAAGCAAGCAGCAG 1155
Qy 1081 CTGAGGAGGACGAGAAAGCACTGTATGTTTGAAGCTGCGGAGGAGCTGGAGGAG 1140
Db 1156 CTGAGGAGGACGAGAAAGCACTGTATGTTTGAAGCTGCGGAGGAGCTGGAGGAG 1215
Qy 1141 TGGCAAGTCGTCGTGGGAAGTTTCTGGCCATCAATGCCACAACATGTCTGTGCTGT 1200
Db 1216 TGGCAAGTCGTCGTGGGAAGTTTCTGGCCATCAATGCCACAACATGTCTGTGCTGT 1275
Qy 1201 CGCCGAGCCCGAGGGCTCTCCCGGCTGCCACTTGGGAGCGGCTTCTGCACCTC 1260
Db 1276 CGCCGAGCCCGAGGGCTCTCCCGGCTGCCACTTGGGAGCGGCTTCTGCACCTC 1335
Qy 1261 ATCTCTATCCGGAAATGCTCCAGGTTCAATTTCTGAGATTCTCATCAGGCACCAAC 1320
Db 1336 ATCTCTATCCGGAAATGCTCCAGGTTCAATTTCTGAGATTCTCATCAGGCACCAAC 1395
Qy 1321 CAGCAGGACCAAGTTTGACTTCACTTTTGTGAAGTTTATCGCGTCAAGAAATCCAGTT 1380
Db 1396 CAGCAGGACCAAGTTTGACTTCACTTTTGTGAAGTTTATCGCGTCAAGAAATCCAGTT 1455
Qy 1381 ACCTCGAAGCAATGAGGATGAGGACAGCGACCTCAAGAGGGGGGGAAGCGCTTT 1440
Db 1456 ACCTCGAAGCAATGAGGATGAGGACAGCGACCTCAAGAGGGGGGGAAGCGCTTT 1515
Qy 1441 GGGCACATTTCCAGCAGCACCCCTCTGCTGCTGCACTGCTCCAAACAGCTCTGGAAC 1500
Db 1516 GGGCACATTTGGAGCAGCACCCCTCTGCTGCTGCACTGCTCCAAACAGCTCTGGAAC 1575
Qy 1501 TCGCAGCGGGAGGCTCTGCACAGCCCTGCCATCGAGGTGAGAGTCCACTGCCAGCTGTT 1560
Db 1576 TCGCAGCGGGAGGCTCTGCACAGCCCTGCCATCGAGGTGAGAGTCCACTGCCAGCTGTT 1635
Qy 1561 CGACTCTTTGACAGGAAATGAGAGAAATCGAAGCCAGACTCACAGCTGA 1614
Db 1636 CGACTCTTTGACAGGAAATGAGAGAAATCGAAGAAATCGAAGCCAGACTCACAGCTGA 1689
```

## RESULT 6

ABA96945

ID ABA96945 standard; cDNA; 4463 BP.

XX AC

XX ABA96945;

XX DT 20-MAY-2002 (first entry)

XX DE Human ceramide kinase hCERK1-encoding cDNA.

XX KW

XX KW Human; ceramide kinase; hCERK1; drug screening; gene therapy;

XX KW neurological disease; inflammation; human immunodeficiency virus;

XX KW HIV infection; type 2 diabetes; obesity; sepsis; arteriosclerosis;

XX KW cancer; neuroprotective; antiinflammatory; anti-HIV; antidiabetic;

XX KW anorectic; antibacterial; antiseptic; antiarteriosclerotic; cytostatic;

XX KW gene; ss.

XX OS

XX Homo sapiens.

XX XX

XX Key Location/Qualifiers

XX FH 124..1737

XX CDS /\*tag= a

XX FT /product= "Human ceramide kinase hCERK1"

XX XX

XX FN WO200196575-A1.

XX PD 20-DEC-2001.

XX XX

XX XX

PF 11-JUN-2001; 2001WO-JF004889.

XX PR 14-JUN-2000; 2000JP-00178039.

XX PA (SANY ) SANKYO CO LTD.

XX PI Sugiyura M, Kono K, Kohama T;

XX DR WPI; 2002-179513/23.

XX DR P-PSDB; AAM49115.

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

Human ceramide kinase gene and the enzyme encoded by it for screening substances as drugs for neurological, inflammatory and other disorders.

Claim 5; Page 46-53; 61pp; Japanese.

This sequence represents cDNA encoding a human ceramide kinase designated hCERK1. The invention relates to hCERK1, nucleic acids encoding it, the expression vectors and host cells containing hCERK1 nucleic acids, the recombinant production of hCERK1 and antibodies specific for hCERK1. The invention also encompasses methods of isolating hCERK1 from samples, the use of hCERK1 in drug screening, and the use of hCERK1 nucleic acid sequences in gene therapy. hCERK1 mediates the ATP-dependent 1-phosphorylation of ceramides and can be used to screen for therapeutic and preventive agents for a wide range of disorders. Such disorders include neurological disease, inflammation, human immunodeficiency virus (HIV) infection, type 2 diabetes, obesity, sepsis, arteriosclerosis and cancer.

Sequence 4463 BP; 1026 A; 1135 C; 1217 G; 1084 T; 0 U; 1 Other;

Query Match 99.6%; Score 1607.6; DB 6; Length 4463;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1610; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGGGGCGCAGCGGGCGGCGGAGCCGCTGCAATCCGCTGTGGTGAAGCAGCAGCGC 60  
Db 124 ATGGGGCGCAGCGGGCGGCGGAGCCGCTGCAATCCGCTGTGGTGAAGCAGCAGCGC 183  
Qy 61 TGGCCGCTGAGCCTGGAGCCCGCGGCGCTCTGCTGCTGGTGGCGAGCCGGGGGCC 120  
Db 184 TGGCCGCTGAGCCTGGAGCCCGCGGCGCTCTGCTGCTGGTGGCGAGCCGGGGGCC 243  
Qy 121 GGAGCCCGCGCCCGCGCGGAGTGCCTGCTGTGCTGTATCTGAGATCATCGCGTT 180  
Db 244 GGAGCCCGCGCCCGCGCGGAGTGCCTGCTGTGCTGTATCTGAGATCATCGCGTT 303  
Qy 181 GAGGAAACAGACGTTTACGGGAAACATCAAGGCGAGTGGAAAAATGGCAGAAAAATG 240  
Db 304 GAGGAAACAGACGTTTACGGGAAACATCAAGGCGAGTGGAAAAATGGCAGAAAAATG 363  
Qy 241 CCTTACGCTTTTACAGTTTACTGTAAAGAGACGACGCGCACCCCTGGAAGTGGCG 300  
Db 364 CCTTACGCTTTTACAGTTTACTGTGTAAAGAGACGACGCGCACCCCTGGAAGTGGCG 423  
Qy 301 CAGGTGACTTTTCTGGTGTCCAGAGGACGAGCTGTGTCACTTGTGGTCCAGACCCCTGCGG 360  
Db 424 CAGGTGACTTTTCTGGTGTCCAGAGGACGAGCTGTGTCACTTGTGGTCCAGACCCCTGCGG 483  
Qy 361 GAGATGCTGGAGAGCTGAGCTCCAGACCAAGCAATTTACTGGTATTTTATCAACCCGTTT 420  
Db 484 GAGATGCTGGAGAGCTGAGCTCCAGACCAAGCAATTTACTGGTATTTTATCAACCCGTTT 543  
Qy 421 GGAGGAAAGGACAAAGCGGATATATGAAGAAAAAGTGGCAACCACTGTTTACACCTTA 480  
Db 544 GGAGGAAAGGACAAAGCGGATATATGAAGAAAAAGTGGCAACCACTGTTTACACCTTA 603  
Qy 481 GCCTCCATCACCCTGACATCATCGTTTACTGAACATGCTAATCAGGCCAAGGAGACTGTG 540  
Db 604 GCCTCCATCACCCTGACATCATCGTTTACTGAACATGCTAATCAGGCCAAGGAGACTGTG 663  
Qy 541 TATGAGATTAACTAGACAAATACGACGGATCGTCTGTGTGGCGGAGATGGTATGTTTC 600

Db 664 TATGAGATTAAATAGACAAATACGACGGCATCTGTGTGTCGGCGGAGATGTTATGTTT 723  
QY 601 ACGGAGTCTGACCGTCTGATTTGGAGAGCGCAGAGAGCGCGGGTTCACACAGAAC 660  
Db 724 ACGGAGTCTGACCGTCTGATTTGGAGAGCGCAGAGAGCGCGGGTTCACACAGAAC 783  
QY 661 CACCCCGGGCTGTGCTGCTCCCTAGCTAGCTCCGATTTGGAATCATTTCCCGCAGGTC 720  
Db 784 CACCCCGGGCTGTGCTGCTCCCTAGCTAGCTCCGATTTGGAATCATTTCCCGCAGGTC 843  
QY 721 ACGGAGTCTGCTGATTTCTTCCACCGTGGCAGCAGCAGAACCTTCGCGCTGCTGAT 780  
Db 844 ACGGAGTCTGCTGATTTCTTCCACCGTGGCAGCAGCAGAACCTTCGCGCTGCTGAT 903  
QY 781 ATCGTGTGGGAGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
Db 904 ATCGTGTGGGAGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 963  
QY 841 CTTGCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
Db 964 CTTGCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1023  
QY 901 GAGAGAAACGGTGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTC 960  
Db 1024 GAGAGAAACGGTGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTC 1083  
QY 961 TCCACCACTGCTATGAGAGGACAGTGTCTTCTCTCTGCAACAACACGCTGGATCT 1020  
Db 1084 TCCACCACTGCTATGAGAGGACAGTGTCTTCTCTCTGCAACAACACGCTGGATCT 1143  
QY 1021 CCAAGGATAGGAAGCCCTGCGGCGAGTCTTGTGTCAGGCAAGCAAGCAGCAG 1080  
Db 1144 CCAAGGATAGGAAGCCCTGCGGCGAGTCTTGTGTCAGGCAAGCAAGCAGCAG 1203  
QY 1081 CTGAGGAGGAGCAGAGAAACACTGTATGTTTGAAGCTTGCAGGAGAGCTGGAGGAG 1140  
Db 1204 CTGAGGAGGAGCAGAGAAACACTGTATGTTTGAAGCTTGCAGGAGAGCTGGAGGAG 1263  
QY 1141 TGGCAAGTCTGTGTGGAGATTTCTGCGCATCAATGCAACAACATGCTGTGCTTGT 1200  
Db 1264 TGGCAAGTCTGTGTGGAGATTTCTGCGCATCAATGCAACAACATGCTGTGCTTGT 1323  
QY 1201 CGCCGAGCCCGAGGGGCTCTCCCGGCTGCGCACTTGGGAGAGGCTTCTGACCTC 1260  
Db 1324 CGCCGAGCCCGAGGGGCTCTCCCGGCTGCGCACTTGGGAGAGGCTTCTGACCTC 1383  
QY 1261 ATCTCTCATCCGAAATGCTCCAGGTTCAATTTCTGAGATTTCTCATGAGCAGACCAAC 1320  
Db 1384 ATCTCTCATCCGAAATGCTCCAGTTCATTTCTGAGATTTCTCATGAGCAGACCAAC 1443  
QY 1321 CAGCAGGACCACTTGAATTTCACTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTT 1380  
Db 1444 CAGCAGGACCACTTGAATTTCACTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTT 1503  
QY 1381 ACGTCGAGCAGCAGGAGTGAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGT 1440  
Db 1504 ACGTCGAGCAGCAGGAGTGAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGT 1563  
QY 1441 GGGCAGATTTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGT 1500  
Db 1564 GGGCAGATTTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGT 1623  
QY 1501 TGCGAGCGGGAGGTCTGTCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGT 1560  
Db 1624 TGCGATGGGAGGTCTGTCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGT 1683  
QY 1561 CGACTCTTTGACGAGGATTTGAAGAGATTCGAGAGCAGCAGCAGCAGCAGCAGTGA 1614  
Db 1684 CGACTCTTTGACGAGGATTTGAAGAGATTCGAGAGCAGCAGCAGCAGCAGCAGTGA 1737

ID ADN62844 standard; DNA; 1740 BP.  
XX AC ADN62844;  
XX XX  
DT 01-JUL-2004 (first entry)  
XX Human NOV9a DNA.  
XX ds; gene; human; NOVX; metabolic disorder; diabetes; obesity;  
KW infectious disease; anorexia; cancer; cancer-associated cachexia;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW immune disorder; haematopoietic disorder; dyslipidaemia;  
KW metabolic syndrome X; wasting disorder.  
XX Homo sapiens.  
XX PN US2004038223-A1.  
XX XX  
XX PD 26-FEB-2004.  
XX PF 01-OCT-2002; 2002US-00262511.  
XX 02-OCT-2001; 2001US-0326483P.  
PR 05-OCT-2001; 2001US-0327435P.  
PR 05-OCT-2001; 2001US-0327449P.  
PR 09-OCT-2001; 2001US-0327917P.  
PR 09-OCT-2001; 2001US-0328029P.  
PR 09-OCT-2001; 2001US-0328044P.  
PR 09-OCT-2001; 2001US-0328056P.  
PR 12-OCT-2001; 2001US-0328849P.  
PR 15-OCT-2001; 2001US-0329414P.  
PR 17-OCT-2001; 2001US-0330142P.  
PR 18-OCT-2001; 2001US-0330309P.  
PR 22-OCT-2001; 2001US-0341058P.  
PR 24-OCT-2001; 2001US-0339266P.  
PR 29-OCT-2001; 2001US-0349575P.  
PR 01-NOV-2001; 2001US-0346357P.  
PR 17-APR-2002; 2002US-0373260P.  
PR 19-APR-2002; 2002US-0373815P.  
PR 19-APR-2002; 2002US-0373817P.  
PR 19-APR-2002; 2002US-0373826P.  
PR 19-APR-2002; 2002US-0373884P.  
PR 22-APR-2002; 2002US-0374977P.  
PR 16-MAY-2002; 2002US-0381037P.  
PR 16-MAY-2002; 2002US-0381038P.  
PR 16-MAY-2002; 2002US-0381042P.  
PR 17-MAY-2002; 2002US-0381642P.  
PR 28-MAY-2002; 2002US-0383656P.  
PR 29-MAY-2002; 2002US-0383831P.  
PR 25-JUN-2002; 2002US-0391335P.  
XX (SMIT/) SMITHSON G.  
PA (MILL/) MILLET I.  
PA (PEYM/) PEYMAN J A.  
PA (KEKU/) KEKUDA R.  
PA (JUJ/) JU J.  
PA (LILL/) LI L.  
PA (GUOX/) GUO X.  
PA (PATT/) PATTURAJAN M.  
PA (SPYT/) SPYTEK K A.  
PA (EDIN/) EDINGER S R.  
PA (ELLE/) ELLERMAN K.  
PA (MALY/) MALYANKAR U M.  
PA (ORTT/) ORT T.  
PA (GORM/) GORMAN L.  
PA (ZERH/) ZERHUSEN B D.  
PA (ANDE/) ANDERSON D W.  
PA (ZHON/) ZHONG M.  
PA (CATT/) CATTERTON E.  
PA (JIWW/) JI W.  
PA (MILL/) MILLER C E.  
PA (RAST/) RASTELLI L.











Db 481 ATCGTCTGTGTCGGCGGAGATGGTATGTTTCAGCGAGGTGCTGCACGGTCTGTGATTGGGAGG 540  
Qy 631 AGCAGAGAGCGCCGGGGTGCACAGAAACCAACCCCGGGCTGTGTCTGCTCCCAAGTACG 690  
Db 541 ACACAGAGAGCGCCGGGGTGCACAGAAACCAACCCCGGGCTGTGTCTGCTCCCAAGTACG 600  
Qy 691 CTCGGATTTGGAATCAATTCCTCCGAGGTCACCGGCTCAACGGACTGGTGTGTACTCCACCGTGGC 750  
Db 601 CTCGGATTTGGAATCAATTCCTCCGAGGTCACCGGCTCAACGGACTGGTGTGTACTCCACCGTGGC 660  
Qy 751 ACCAGCGACGCAAAACCTTCGGCGCTGCATATCGTTGTTGGGACTCGCTGGCCATGGAT 810  
Db 661 ACCAGCGACGCAAAACCTTCGGCGCTGCATATCGTTGTTGGGACTCGCTGGCCATGGAT 720  
Qy 811 GTGTCTCTAGTCACACAAAGCACACTCTCTGCTACTCCGTGTCTCTGCTGGGCTAC 870  
Db 721 GTGTCTCTAGTCACACAAAGCACACTCTCTGCTACTCCGTGTCTCTGCTGGGCTAC 780  
Qy 871 GGTCTTCTACGGGACATCATCAAGGACAGTGAAGAAACGGTGGTCTGCTCCAGA 930  
Db 781 GGTCTTCTACGGGACATCATCAAGGACAGTGAAGAAACGGTGGTCTGCTCCAGA 840  
Qy 931 TAGGACTTTTCAAGTTTAAAGACCTTCTCTCTCCCACTGCTATGAAGGACAGTGTCC 990  
Db 841 TAGGACTTTTCAAGTTTAAAGACCTTCTCTCTCCCACTGCTATGAAGGACAGTGTCC 900  
Qy 991 TTCTCTCTCTGCAACACACAGTGGATCTCAAGGATAGAAAGCCCTGCGGACAGA 1050  
Db 901 TTCTCTCTCTGCAACACACAGTGGATCTCAAGGATAGAAAGCCCTGCGGACAGA 960  
Qy 1051 TGCTTTGTTTTCAGGCAAGGACAGCAGCTGGAGGAGGACAGAAAGCACATGTAT 1110  
Db 961 TGCTTTGTTTTCAGGCAAGGACAGCAGCTGGAGGAGGACAGAAAGCACATGTAT 1020  
Qy 1111 GGTTTGGAAGCTCGGAGACGTGGAGGAGTGGCAAGTCTGTGGGAAGTTTCTGGCC 1170  
Db 1021 GGTTTGGAAGCTCGGAGACGTGGAGGAGTGGCAAGTCTGTGGGAAGTTTCTGGCC 1080  
Qy 1171 ATCAATGCAAAACATGCTCTGTCTGTCGCGGAGCCCGAGGGCTCTCCCGGT 1230  
Db 1081 ATCAATGCAAAACATGCTCTGTCTGTCGCGGAGCCCGAGGGCTCTCCCGGT 1140  
Qy 1231 GCCACTTGGGAGACGGTCTTCTGACCTCATCTCATCGGAAATGCTCCAGGTTCAAT 1290  
Db 1141 GCCCACTTGGGAGACGGTCTTCTGACCTCATCTCATCGGAAATGCTCCAGGTTCAAT 1200  
Qy 1291 TTTCTGAGATTTCTCATCAGGCACACCAACAGCAGGACAGTTTGACTTCACTTTTGT 1350  
Db 1201 TTTCTGAGATTTCTCATCAGGCACACCAACAGCAGGACAGTTTGACTTCACTTTGT 1260  
Qy 1351 GAAGTTTATCGCTCAAGAAATTCAGTTTACGTTCGAAGCACATGGAGGATGAGGACAGC 1410  
Db 1261 GAAGTTTATCGCTCAAGAAATTCAGTTTACGTTCGAAGCACATGGAGGATGAGGACAGC 1320  
Qy 1411 GACCTCAGGAGGGGGGAGAGCGCTTTGGGCAATTTGACAGACGACCCCTCTGTC 1470  
Db 1321 GACCTCAGGAGGGGGGAGAGCGCTTTGGGCAATTTGACAGACGACCCCTCTGTC 1380  
Qy 1471 TGCTGACCGCTTCCAAAGCTCTTGGAACTCGCAGCGGGAGTCTTGACAGCCCTGCC 1530  
Db 1381 TGCTGACCGCTTCCAAAGCTCTTGGAACTCGCAGCGGGAGTCTTGACAGCCCTGCC 1440  
Qy 1531 ATCGAGGTGAGAGTCCACTGCGAGCTGGTTGACTCTTTTGACAGGAGAAATGAAGAGAT 1590  
Db 1441 ATCGAGGTGAGAGTCCACTGCGAGCTGGTTGACTCTTTTGACAGGAGAAATGAAGAGAT 1500  
Qy 1591 CCGAAGCCAGACTCACACAGCTGA 1614  
Db 1501 CCGAAGCCAGACTCACACAGCTGA 1524

RESULT 10  
ACC70838

ID ACC70838 standard; DNA; 3975 BP.  
XX AC ACC70838;  
XX AC ACC70838;  
DT 20-NOV-2003 (first entry)  
XX Human Sphingosine kinase 4 coding sequence.  
XX Human; enzyme; haemostatic; sphingosine kinase 4; SPHK4;  
KW platelet transfusion; platelet stabiliser; gene; ds.  
XX Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 1..1251  
FT /\*tag= a  
FT /product= "Sphingosine kinase 4"  
XX WO2003031627-A1.  
XX PD 17-APR-2003.  
XX 28-SEP-2001; 2001WO-JP008537.  
XX 28-SEP-2001; 2001WO-JP008537.  
XX (HOKK-) HOKKAIDO TECHNOLOGY LICENSING OFFICE CO.  
XX (CHBI-) CHEM BIOLOGY INST.  
XX Igarashi Y, Kihara A;  
XX WPI; 2003-354917/33.  
DR P-PSDB; ABR56301.  
XX Platelet derived polypeptides with sphingosine kinase activity for  
PT treatment of sphingosine related disorders.  
XX Claim 4; Page 30-31; 39pp; Japanese.  
XX The present sequence is the coding sequence for human sphingosine kinase  
CC 4 (SPHK4). The kinase can be used for the diagnosis and treatment of  
CC sphingosine related disorders. The kinase can also be potentially used  
CC for controlling toxicity of platelet transfusion and as a platelet  
CC stabiliser  
XX SQ Sequence 3975 BP; 943 A; 978 C; 1042 G; 1012 T; 0 U; 0 Other;  
Query Match 77.5%; Score 1251; DB 8; Length 3975;  
Best Local Similarity 100.0%; Pred. No. 2.8e-314;  
Matches 1251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 364 ATGCTGGAGAAAGCTGACGTCCAGACCAGCAATTTACTGGTATTTATCAACCCGTTTGA 423  
Db 1 ATGCTGGAGAAAGCTGACGTCCAGACCAGCAATTTACTGGTATTTATCAACCCGTTTGA 60  
Qy 424 GGAAGAGCAAGGACGCGGATATATGAAAGAAAGTGGACCACTGTTTCACTTAGCC 483  
Db 61 GGAAGAGCAAGGACGCGGATATATGAAAGAAAGTGGACCACTGTTTCACTTAGCC 120  
Qy 484 TCCATCACCACTGACATCATCTGTTACTGAAATGCTATATCAGGCCAGGAGACTCTGTAT 543  
Db 121 TCCATCACCACTGACATCATCTGTTACTGAAATGCTATATCAGGCCAGGAGACTCTGTAT 180  
Qy 544 GAGATTACATAGACAAATACGACGCGCATCTGTGTGTCGGCGGAGATGGTATGTTACG 603  
Db 181 GAGATTACATAGACAAATACGACGCGCATCTGTGTGTCGGCGGAGATGGTATGTTACG 240  
Qy 604 GAGGTGTGACGGTCTGATTTGGAGGACGACAGAGGAGCGCCGGGGTTCAGACAGAACAC 663  
Db 241 GAGGTGTGACGGTCTGATTTGGAGGACGACAGAGGAGCGCCGGGGTTCAGACAGAACAC 300  
Qy 664 CCCGGGGTGTGTGGTCTCCAGTAGCCTCCGGATTGGAATCATTCGCCGAGGGTCAACG 723

Db 301 CCCGGGCTGTGCTGCTCCAGTAGCTCCGGAATGGAATCATTCOCGAGGCTCAACG 360  
Qy 724 GACTGCGTGTGTACTCCACCGTGGGACACGACGACGAGAAACCTCGCGCTGTATATC 783  
Db 361 GACTGCGTGTGTACTCCACCGTGGGACACGACGACGAGAAACCTCGCGCTGTATATC 420  
Qy 784 GTTGTGGGAGCTCGCTGGCCATGATGTCTCTCAGTCCACCAACAGACACACTCCTT 843  
Db 421 GTTGTGGGAGCTCGCTGGCCATGATGTCTCTCAGTCCACCAACAGACACACTCCTT 480  
Qy 844 CGCTACTCCGTCTCCCTGCTGGCTACGGCTTCTACGGGGACATCATCAAGGACAGTGAG 903  
Db 481 CGCTACTCCGTCTCCCTGCTGGCTACGGCTTCTACGGGGACATCATCAAGGACAGTGAG 540  
Qy 904 AAGAAACCGTGTGGGTCTTGGCCAGATACGACTTTTTCAGGTTTTAAAGACCTTCTCTCC 963  
Db 541 AAGAAACCGTGTGGGTCTTGGCCAGATACGACTTTTTCAGGTTTTAAAGACCTTCTCTCC 600  
Qy 964 CACCACCTGCTATGAGGACAGTGTCTTCTCCCTGCAACACACACGGTGGGATCTCCA 1023  
Db 601 CACCACCTGCTATGAGGACAGTGTCTTCTCCCTGCAACACACACGGTGGGATCTCCA 660  
Qy 1024 AGGATAGGAGCCCTGCGGGCAGGATGCTTTGTTTTCAGGCAAGCAAGCAGCAGCTG 1083  
Db 661 AGGATAGGAGCCCTGCGGGCAGGATGCTTTGTTTTCAGGCAAGCAAGCAGCAGCTG 720  
Qy 1084 GAGGAGGACAGAGAAACACTGTATGTTTGGAGCTGCGGAGGAGCTGGAGAGTGG 1143  
Db 721 GAGGAGGACAGAGAAACACTGTATGTTTGGAGCTGCGGAGGAGCTGGAGAGTGG 780  
Qy 1144 CAAGTCGTCTGTGGAGTTTCTGSCCATCAATGCCAACAACATGTCTGTGCTGTGTCG 1203  
Db 781 CAAGTCGTCTGTGGAGTTTCTGSCCATCAATGCCAACAACATGTCTGTGCTGTGTCG 840  
Qy 1204 CGAGCCCCAGGGGCTCTCCCGGCTGCCACATTTGGAGAGCGGTCTTCGACCTATC 1263  
Db 841 CGAGCCCCAGGGGCTCTCCCGGCTGCCACATTTGGAGAGCGGTCTTCGACCTATC 900  
Qy 1264 CTCATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTCTCATCAGGACACCAACCGAG 1323  
Db 901 CTCATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTCTCATCAGGACACCAACCGAG 960  
Qy 1324 CAGGACCAAGTTGACTTCACTTTTGTGAAGTTTATCCGCTCAAGAAATTCAGTTTACG 1383  
Db 961 CAGGACCAAGTTTGAATCTTCACTTTTGTGAAGTTTATCCGCTCAAGAAATTCAGTTTACG 1020  
Qy 1384 TCGAAGCACATGAGGATGAGGACAGCGACCTCAAGGAGGGGGGAGAAAGCGCTTTGGG 1443  
Db 1021 TCGAAGCACATGAGGATGAGGACAGCGACCTCAAGGAGGGGGGAGAAAGCGCTTTGGG 1080  
Qy 1444 CACATTTGACGAGCCACCCCTCTGCTGCTGACCGCTCTCAACAGCTCTCTGGAACCTGC 1503  
Db 1081 CACATTTGACGAGCCACCCCTCTGCTGCTGACCGCTCTCAACAGCTCTCTGGAACCTGC 1140  
Qy 1504 GACGGGAGGTCCTGSCACAGCCCTGCCATCAGAGTTCAGAGTCCAATGCCAGCTGGTTCCA 1563  
Db 1141 GACGGGAGGTCCTGSCACAGCCCTGCCATCAGAGTTCAGAGTCCAATGCCAGCTGGTTCCA 1200  
Qy 1564 CTCCTTTCACGAGGATTTGAAGAGATTCGAAGCCAGACTCACACAGCTGA 1614  
Db 1201 CTCCTTTCACGAGGATTTGAAGAGATTCGAAGCCAGACTCACACAGCTGA 1251

RESULT 11  
ID ADS10370 standard; DNA; 4702 BP.  
XX AC ADS10370;  
XX DT 16-DEC-2004 (first entry)  
XX DE Human therapeutic DNA - SEQ ID 607.  
XX

KW antiinflammatory; neuroprotective; antianaemic; cytostatic; vulneryary;  
KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;  
XX aplastic anaemia; cancer; wound healing; gene therapy; ds; gene.  
OS Homo sapiens.  
XX WO2004080148-A2.  
PN 23-SEP-2004.  
XX 30-SEP-2003; 2003WO-US030720.  
XX 02-OCT-2002; 2002US-0416186P.  
PR (NUVE-) NUVELO INC.  
XX Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y; Zhou P;  
PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AU, Weng G, Zhou P;  
XX WPI; 2004-668857/65.  
DR P-PSDB; ADS11054.  
XX New polynucleotide, useful in preparing a composition for diagnosing or  
PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,  
PT aplastic anemia or cancer for promoting wound healing.  
XX Claim 1; SEQ ID NO 607; 718pp; English.  
XX The invention relates to a novel isolated polynucleotide and the encoded  
CC polypeptide. The molecules of the invention demonstrate antiinflammatory,  
CC neuroprotective, antianaemic, cytostatic and vulneryary activities and may  
CC be useful in preparing a composition for diagnosing or treating  
CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell  
CC disorders, such as aplastic anaemia or cancer, as well as for promoting  
CC wound healing. The molecules may also be utilised during gene therapy  
CC procedures. The current sequence is that of a human therapeutic DNA of  
CC the invention. The current sequence is not shown explicitly within the  
CC specification but can be accessed from the WIPO web-site.  
XX Sequence 4702 BP; 970 A; 1226 C; 1364 G; 1118 T; 0 U; 24 Other;  
SQ

Query Match 63.0%; Score 1017.6; DB 13; Length 4702;  
Best Local Similarity 81.4%; Pred. No. 1e-253;  
Matches 1261; Conservative 0; Mismatches 214; Indels 75; Gaps 4;  
Qy 140 CGGATGCTGCTGTGCTGTATCTGAGATCATCGCGTTGAGAAACACAGCTTCACG 199  
Db 1394 CAGATGCTGCTGTGCTGTATCTGAGATCATCGCGTTGAGAAACACAGCTTCACG 1453  
Qy 200 GGAAACATCAAGGACAGTGGAAATGGCAGAAATGGAAAGCCTTACGTTTTCAGTTC 259  
Db 1454 GGAAACATCAAGGACAGTGGAAATGGCAGAAATGGAAAGCCTTACGTTTTCAGTTC 1513  
Qy 260 ACTGTGTAAGAGAGACGACGCGCACCGCTGGAAAGTGGCGCAGTCACTTCTGGTGTG 319  
Db 1514 ACTGTGTAAGAGAGACGACGCGCACCGCTGGAAAGTGGCGCAGTCACTTCTGGTGTG 1573  
Qy 320 CAGAGGACGAGCTGTGTCACCTGTGGCTGACAGCCCTGCGGAGATGCTGAGAGCTGA 379  
Db 1574 CAGAGGACGAGCTGTGTCACCTGTGGCTGACAGCCCTGCGGAGATGCTGAGAGCTGA 1633  
Qy 380 CGTCCAGAC-----CAAAGCATTTTACTGTTATTTTATCAACCCGTTTGGAGAA 427  
Db 1634 TTACTGAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAACATAGACAAATACG 1693  
Qy 428 AAGGACAAAGCAAGCGGATATATGAAAGAAAGTGGCACCACTGTTACCTTAGCCTCCA 487  
Db 1694 ACGGCATGCTGTGTGCGCGGAGATGTTATGTTTCAGCGAGTGTGTCACCGTCTGATG 1753  
Qy 488 TCACCACTGACATCATGTTTACTGAAACATGCTAATCAGGCCCAAGGAGACTCTGTATGAG 547  
Db 1754 GGAGGACGACGAGGCGCGGGTTCAGCCAGAACCCCGCGGCTGTGTGTTGCCCA 1813

QY 548 TTAACAT-----AGCAAAATACGACGGCATCGTCTGTGTGCGGGAGATG----- 592  
Db 1814 GTAGCTCCGGATTGGAATATTCGCGCAGGCTTTGCAAAATGACCTGGCGGGAGGAGG 1873  
QY 593 -GTATGTTGAGGAGGTGTGCGAGGTCTGATTTGGAGGACGCGAGAGCCCGGGGTC 651  
Db 1874 TGTCTGTCTCTCTGCGCCCTGTGTCTGGCCCGGAGGGTGGCGCATGGTGCACTTTCA 1933  
QY 652 GACCAAGAACCCCGGCTGTGTGTCCTCCAGTAGCTCGGATCGGAATCATTCCTCC 711  
Db 1934 CTCTGTGACAGTCTGGGATGTGAGACCGCAGTCATCCCATTTTATGATGAAGACA 1993  
QY 712 GCA-----GGGTCAACGG 724  
Db 1994 GGAGACTGGGAGCATGTGGGCCCCGGTGAGAACGCTGTGTGGCTTGGAGGGTCAACGG 2053  
QY 725 ACTCGGTGTGTTATCTCCACCGTGGGACCGAGGACGAGAAACCTCGGCGCTGCATATCG 784  
Db 2054 ACTCGGTGTGTTACTCCACCGTGGGACCGAGGACGAGAAACCTCGGCGCTGCATATCG 2113  
QY 785 TTGTTGGGACTCGCTGGCCATGATGTGTCTCTAGTCCACCAACAGCACACTCCTTC 844  
Db 2114 TTGTTGGGACTCGCTGGCCATGATGTGTCTCTAGTCCACCAACAGCACACTCCTTC 2173  
QY 845 GCTACTCTCGCTGTCCCTGTGGGCTACGGCTTCTAGGGGACATCATCAAGGACAGTGAGA 904  
Db 2174 GCTACTCTCGCTGTCCCTGTGGGCTACGGCTTCTAGGGGACATCATCAAGGACAGTGAGA 2233  
QY 905 AGAAACGGTGTGGGTCTTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCC 964  
Db 2234 AGAAACGGTGTGGGTCTTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCC 2293  
QY 965 ACCACTGCTATGAAGGACAGTGTCTCTCTCCCTGCGCAACACAGGTTGGATCTCAA 1024  
Db 2294 ACCACTGCTATGAAGGACAGTGTCTCTCTCCCTGCGCAACACAGGTTGGATCTCAA 2353  
QY 1025 GGGATAGGAAGCCCTCGCGGCGAGATGCTTTGTTTGCAGGCAAGAGCAGCAGCTGG 1084  
Db 2354 GGGATAGGAGCCCTCGCGGCGAGATGCTTTGTTTGCAGGCAAGAGCAGCAGCTGG 2413  
QY 1085 AGGAGGACGAGAAGAACACCTGTATGTTTGGAAAGCTCGGAGGACGTGGAGGATGGC 1144  
Db 2414 AGGAGGACGAGAAGAACACCTGTATGTTTGGAAAGCTCGGAGGACGTGGAGGATGGC 2473  
QY 1145 AAGTCGTCTGTGGGAGTTTCTGGGCATCAATGCCAACAATGCTCTGTGTCTGTGCGC 1204  
Db 2474 AAGTCGTCTGTGGGAGTTTCTGGGCATCAATGCCAACAATGCTCTGTGTCTGTGCGC 2533  
QY 1205 GGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGCGGTCTTCTGACCTCATCC 1264  
Db 2534 GGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGCGGTCTTCTGACCTCATCC 2593  
QY 1265 TCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAACCCAGC 1324  
Db 2594 TCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAACCCAGC 2653  
QY 1325 AGGACAGTTTCACTTCACTTTTGTGAATTTATCGGTCGAAGAAATTCAGATTACGT 1384  
Db 2654 AGGACAGTTTCACTTCACTTTTGTGAATTTATCGGTCGAAGAAATTCAGATTACGT 2713  
QY 1385 CGAAGCACATGAGGATGAGGACGACCTCAAGAGGGGGGGAAGAGCGCTTTGGGC 1444  
Db 2714 CGAAGCACATGAGGATGAGGACGACCTCAAGAGGGGGGGAAGAGCGCTTTGGGC 2773  
QY 1445 ACATTTGAGGACGACCCCTCTCTGCTGTGACCGTCTCCAAACAGCTCTCTGGAACCTGCG 1504  
Db 2774 ACATTTGAGGACGACCCCTCTCTGCTGTGACCGTCTCCAAACAGCTCTCTGGAACCTGCG 2833  
QY 1505 ACGGGAGGTCTGTGACACCCCTGCCATCGAGGTGAGGTCCATGCGCAGCTGTGTCGAC 1564  
Db 2834 ACGGGAGGTCTGTGACACCCCTGCCATCGAGGTGAGGTCCATGCGCAGCTGTGTCGAC 2893  
QY 1565 TCTTTTGACGAGGAATTGAAGAGATCCGAGCCAGACTCAGACAGCTGA 1614

Db 2894 TCTTTCAGAGGAATTGAAGAGATCCGAGCCAGACTCAGACAGCTGA 2943

RESULT 12  
AAS77728

ID AAS77728 standard; cDNA; 2241 BP.

XX AAS77728;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #13532.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO2001175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG13541.

PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

XX Claim 1; SEQ ID NO 13532; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 2241 BP; 453 A; 627 C; 700 G; 461 T; 0 U; 0 Other;

..Query Match 57.7%; Score 931; DB 5; Length 2241;  
Best Local Similarity 80.2%; Pred. No. 2.4e-231;  
Matches 1273; Conservative 0; Mismatches 5; Indels 309; Gaps 3;

QY 140 CGGATGCTGCTGTGCTGCTATCTGAGATCATCGCGTTTGAGGAAACAGCGTTTCAGG 199

Db 659 CAGATGCTGCTGTGCTGCTATCTGAGATCATCGCGTTTGAGGAAACAGCGTTTCAGG 718

Qy 200 GGAAACATCAAGGCAGTGGAAAAATGGCAGAAAAATGGAAAAAGCCTTACGCTTTTACAGTTC 259  
Db |||||  
Qy 719 GGAAACATCAAGGCAGTGGAAAAATGGCAGAAAAATGGAAAAAGCCTTACGCTTTTACAGTTC 778  
Db |||||  
Qy 260 ACTGTGTAAAGAGACGACGCGCACCGCTGGAGTGGCGCAGGTGACTTCTTCTGGTGTG 319  
Db |||||  
Qy 779 ACTGTGTAAAGAGACGACGCGCACCGCTGGAGTGGCGCAGGTGACTTCTTCTGGTGTG 838  
Db |||||  
Qy 320 CAGAGGACGCTGTGTCACTTGTGGCTGCAGACCCCTGCGGAGATGCTGGAGAAGCTGA 379  
Db |||||  
Qy 839 CAGAGGACGCTGTGTCACTTGTGGCTGCAGACCCCTGCGGAGATGCTGGAGAAGCTGA 898  
Db |||||  
Qy 380 CGTCAGACCAAGCATTTACTGTGTATTTATCAACCCGTTTGGAGAAAAGGACAAAGCA 439  
Db |||||  
Qy 899 ----- 898  
Qy 440 AGCGGATATATGAAGAAAAGTGGCACCATCTGTTTCACTTAGCTTCCATCACCACCTGACA 499  
Db |||||  
Qy 899 ----- 898  
Qy 500 TCATCGTTACTGAACATCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACA 559  
Db |||||  
Qy 899 -----TTACTGAACATCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACA 952  
Db |||||  
Qy 560 AATACGACGGCATCGTCTGTGTGCGCGGAGATGATGTTTCAAGCAGAGTGCTGCACCGTTC 619  
Db |||||  
Qy 953 AATACGACGGCATCGTCTGTGTGCGCGGAGATGATGTTTCAAGCAGAGTGCTGCACCGTTC 1012  
Db |||||  
Qy 620 TGATTGGAGACGACGAGAGAGCGCGGGTGCAGACCAACACCCCGGCTGTGCTGG 679  
Db |||||  
Qy 1013 TGATTGGAGAGACGACGAGAGAGCGCGGGTGCAGACCAACACCCCGGCTGTGCTGG 1072  
Db |||||  
Qy 680 TCCCGAGTAGCTCCGGATTGGAATCATTTCCCGCA----- 714  
Db |||||  
Qy 1073 TCCCGAGTAGCTCCGGATTGGAATCATTTCCCGCAGAGCATGTGGCCCCCGGTGAGAACG 1132  
Db |||||  
Qy 715 -----GGGTCAACGGACTGCGTGTGTTACTCCACCGTGGGACCCAGCGACG 760  
Db |||||  
Qy 1133 CTGTGTGGCTTGACGGGTCAACGGACTGCGTGTGTTACTCCACCGTGGGACCCAGCGACG 1192  
Db |||||  
Qy 761 CAGAAACCTCGCGCTGCATATCGTTGTT----- 789  
Db |||||  
Qy 1193 CAGAAACCTCGCGCTGCATATCGTTGTTGCTGTGCTGCCCGAGGCTCGGAAACCAACCCG 1252  
Db |||||  
Qy 790 ----- 789  
Db |||||  
Qy 1253 CATCCCGCATACTGCTGTGGCAGTGGGACGGGACGCTGTGCTGGGCTGCCAGCTGT 1312  
Db |||||  
Qy 790 -----GGGGACT 796  
Db |||||  
Qy 1313 GGAAGCATGCTCTGTGAGGCTCGAGGCTTCAGTCCAGGATGCAGAGCCCCGGGACT 1372  
Db |||||  
Qy 797 CGCTGGCCATGGATGTGTCCTCAGTCCACACCAAGCAGACATCTCTTCTGCTACTCCGTTGT 856  
Db |||||  
Qy 1373 CGCTGGCCATGGATGTGTCCTCAGTCCACCAAGCAGACATCTCTTCTGCTACTCCGTTGT 1432  
Db |||||  
Qy 857 CCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGCAGAGAAACGTTGTT 916  
Db |||||  
Qy 1433 CCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGCAGAGAAACGTTGTT 1492  
Db |||||  
Qy 917 TGGGTCTTTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCTCCACCACTGCTATG 976  
Db |||||  
Qy 1493 TGGGTCTTTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCTCCACCACTGCTATG 1552  
Db |||||  
Qy 977 AAGGACAGTGTCTTCTCCCTGCACAAACACAGCGTGGATCTTCCAAAGGATAGGAAGC 1036  
Db |||||  
Qy 1553 AAGGACAGTGTCTTCTCCCTGCACAAACACAGCGTGGATCTTCCAAAGGATAGGAAGC 1612  
Db |||||  
Qy 1037 CCTGCGGGCAGGATGCTTGTGTCAGGCAAGCAAGCAGCAGCTGGAGGAGGACAGA 1096  
Db |||||  
Qy 1613 CCTGCGGGCAGGATGCTTGTGTCAGGCAAGCAAGCAGCAGCTGGAGGAGGACAGA 1672  
Db |||||

Qy 1097 AGAAAGCACTGTATGTTTGGAAAGCTCGGAGGAGCTGGAGGAGTGGCAAGTCTCTGTG 1156  
Db |||||  
Qy 1673 AGAAAGCACTGTATGTTTGGAAAGCTCGGAGGAGCTGGAGGAGTGGCAAGTCTCTGTG 1732  
Db |||||  
Qy 1157 GGAAGTTTCTGGCCATCAATGCCACAAACATGTCCTGTGCTGTGTCGCGGAGCCCCCAGGG 1216  
Db |||||  
Qy 1733 GGAAGTTTCTGGCCATCAATGCCACAAACATGTCCTGTGCTGTGTCGCGGAGCCCCCAGGG 1792  
Db |||||  
Qy 1217 GCCTCTCCCGGGCTGCCACTTGGGAGACGGGTCTTCTGACCTCATCTCATCCGGAAT 1276  
Db |||||  
Qy 1793 GCCTCTCCCGGGCTGCCACTTGGGAGACGGGTCTTCTGACCTCATCTCATCCGGAAT 1852  
Db |||||  
Qy 1277 GCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGACACCAACCCAGCAGACCCAGTTTG 1336  
Db |||||  
Qy 1853 GCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGACACCAACCCAGCAGACCCAGTTTG 1912  
Db |||||  
Qy 1337 ACTTCACCTTTGTTGAAGTTTATCGCGTCAAGAAATTCAGTTTACGTTTACGTTGAGCACATGG 1396  
Db |||||  
Qy 1913 ACTTCACCTTTGTTGAAGTTTATCGCGTCAAGAAATTCCTGTTTACGTTGAGCACATGG 1972  
Db |||||  
Qy 1397 AGGATGAGGACAGCGACCTCAAGGAGGGGGGGAAGAGCGCTTTGGGCACATTTTGCAGCA 1456  
Db |||||  
Qy 1973 AGGATGAGGACAGCGACCTCAAGGAGGGGGGGAAGAGCGCTTTGGGCACATTTTGCAGCA 2032  
Db |||||  
Qy 1457 GCACACCTCTGCTGTGTCACCGTCTCBAACAGCTCCTGGAACCTGCGAGGGAGGTCC 1516  
Db |||||  
Qy 2033 GCCACCCCTCTGCTGTGTCACCGTCTCBAACAGCTCCTGGAACCTGCGAGGGAGGTCC 2092  
Db |||||  
Qy 1517 TGACACGCTTGCCTCGAGTGCAGAG 1543  
Db |||||  
Qy 2093 TGCACAGCTTGCCTCGAGTGCAGAG 2119  
Db |||||  
RESULT 13  
ABL40822  
ID ABL40822 standard; cDNA; 979 BP.  
XX  
XX ABL40822;  
XX AC  
XX AC  
DT 03-JUL-2002 (first entry)  
XX  
DE Human sphingosine kinase-like protein encoding cDNA.  
XX  
XX Human sphingosine kinase-like protein; intracellular signalling; gene;  
KW cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;  
KW autoimmune disease; rheumatoid arthritis; Parkinson's disease; ss.  
XX  
OS Homo sapiens.  
XX  
Key Location/Qualifiers  
FH 2..789  
FT /\*tag= a  
FT /product= "sphingosine kinase-like protein"  
FT /note= "start and stop codons are not indicated"  
XX  
XX WO200228906-A2.  
XX  
PD 11-APR-2002.  
XX  
XX 05-OCT-2001; 2001WO-EP011516.  
XX  
XX 06-OCT-2000; 2000US-0238005P.  
PR 23-AUG-2001; 2001US-0314113P.  
XX  
XX (FARB ) BAYER AG.  
XX  
XX Kossida S, Encinas J;  
XX WPI; 2002-340094/37.  
XX P-PSDB; ABB07854.  
XX  
XX New reagent for modulating the activity of sphingosine kinase-like  
FT protein polypeptide or polynucleotide and treating cancer, asthma,



PT allergy, an autoimmune disease, or a central or peripheral nervous system  
PT disorder.

PS Claim 1; Fig 1; 120pp; English.

The invention relates to a human sphingosine kinase-like protein. The polypeptide can be expressed by standard recombinant methodology. The sphingosine kinase-like protein and gene can be used to regulate intracellular signalling and consequently cell proliferation and apoptosis. Such regulation is useful for treating cancer, allergies (e.g. asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and peripheral nervous system disorders (e.g. Parkinson's disease). The present sequence represents the human sphingosine kinase-like protein encoding cDNA

Sequence 979 BP; 237 A; 259 C; 271 G; 212 T; 0 U; 0 Other;

Query Match	55.9%;	Score 903;	DB 6;	Length 979;
Best Local Similarity	96.6%;	Pred. No. 3.2e-224;		
Matches 946;	Conservative 0;	Mismatches 33;	Indels 33;	Gaps 1;
Qy	387	ACCAAAAGCATTACTGCTATTATTCACACCCGTTTGGAGGAAAGGACAAAGCAAGCCGAT	446	
Db	1	ACCAAAAGCATTACTGCTATTATTCACACCCGTTTGGAGGAAAGGACAAAGCAAGCCGAT	60	
Qy	447	ATATGAAGAAAGTGGCCACACCTGTTTCACCTTAGCCTCCATCACCCTGACATCATCG-	505	
Db	61	ATATGAAGAAAGTGGCCACACCTGTTTCACCTTAGCCTCCATCACCCTGACATCATCG	120	
Qy	506	-----TTACTGAACATGCTTAATCAGGCCAAGGA	533	
Db	121	TAACAAAATCTATGTTAACTATGTAGAGTAATTACTGAACATGCTTAATCAGGCCAAGGA	180	
Qy	534	GACTCTCTATGAGATTAACTAGACAAATACGACGGCATGCTGTCGTGCGCGGAGATGG	593	
Db	181	GACTCTGTATGAGATTAACTAGACAAATACGACGGCATGCTGTCGTGCGCGGAGATGG	240	
Qy	594	TATGTTTCAGCGAGGTGCTGCA CGGTCTCATTTGGAGGACGCGAGGAGCGCGCGGGTCGA	653	
Db	241	TATGTTTCAGCGAGGTGCTGCA CGGTCTCATTTGGAGGACGCGAGGAGCGCGCGGGTCGA	300	
Qy	654	CCAGAACCCCGGGCTGTGCTGTCGCCAGTAGCTCCGGATTCGGAAATCATTCCTCCG	713	
Db	301	CCAGAACCCCGGGCTGTGCTGTCGCCAGTAGCTCCGGATTCGGAAATCATTCCTCCG	360	
Qy	714	AGGTCACACGAGCTCGCTGTGTACTCCACCGTGGGACACACGACGCGAGAAAACCTCGGC	773	
Db	361	AGGTCACACGAGCTCGCTGTGTACTCCACCGTGGGACACACGACGCGAGAAAACCTCGGC	420	
Qy	774	GCTGCATATCGTTTGGGAGCTCGCTGGCCATGGATGTGTCCTCAGTCCACCAACAAG	833	
Db	421	GCTGCATATCGTTTGGGAGCTCGCTGGCCATGGATGTGTCCTCAGTCCACCAACAAG	480	
Qy	834	CACACTCCTTCGCTACTCCGTGTCCTCTGCTGGCTACGGCTTCTACCGGGACATCATCAA	893	
Db	481	CACACTCCTTCGCTACTCCGTGTCCTCTGCTGGCTTCTACCGGGACATCATCAA	540	
Qy	894	GGCAGTGAAGAAACCGTGGTTGGGTCTTGCCAGATACGACTTTTTCAGGTTTTAAAGAC	953	
Db	541	GGCAGTGAAGAAACCGTGGTTGGGTCTTGCCAGATACGACTTTTTCAGGTTTTAAAGAC	600	
Qy	954	CTTCTCTCCACACCTGCTATGAAGGGAAGTGTCTTCTCTCCTGTGCAAAACACGGT	1013	
Db	601	CTTCTCTCCACACCTGCTATGAAGGGAAGTGTCTTCTCTCCTGTGCAAAACACGGT	660	
Qy	1014	GGGATCTCCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTCTGTTTTCAGGCGAAAGCAA	1073	
Db	661	GGGATCTCCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTCTGTTTTCAGGCGAAAGCAA	720	
Qy	1074	GCAGCAGCTGGAGGAGCAGAAAGCACTGTATGTGGTTTGAAGCTGCGGAGGACGT	1133	
Db	721	GCAGCAGCTGGAGGAGCAGAAAGCACTGTATGTGGTTTGAAGCTGCGGAGGACGT	780	



CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 2186 BP; 455 A; 637 C; 669 G; 425 T; 0 U; 0 Other;  
  
Query Match 46.7%; Score 753.8; DB 5; Length 2186;  
Best Local Similarity 99.7%; Pred. No. 2.9e-185;  
Matches 755; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 790 GGGGACTCGCTGGCCATGGATGTCTCTCAGTCACACACACACACACTCTTCGGTAC 849  
Db 876 GGGGACTCGCTGGCCATGGATGTCTCTCAGTCACACACACACACACTCTTCGGTAC 935  
  
Qy 850 TCCGTGTCCCTGCTGGGTACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAA 909  
Db 936 TCCGTGTCCCTGCTGGGTACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAA 995  
  
Qy 910 CGGTGGTGGGTCTTGGCAGATACGACTTTTCAGGTTTAAAGACTTCTCTCCACAC 969  
Db 996 CGGTGGTGGGTCTTGGCAGATACGACTTTTCAGGTTTAAAGACTTCTCTCCACAC 1055  
  
Qy 970 TGCTATGAGGACAGTGTCTCTCTCTCCACACACACAGGTGGATCTCCAGGGAT 1029  
Db 1056 TGCTATGAGGACAGTGTCTCTCTCTCCACACACACAGGTGGATCTCCAGGGAT 1115  
  
Qy 1030 AGGAAGCCCTCCGGGAGGATGCTTTGTTGACGGAAGCAAGCAGCAGCTGGAGG 1089  
Db 1116 AGGAAGCCCTCCGGGAGGATGCTTTGTTGACGGAAGCAAGCAGCAGCTGGAGG 1175  
  
Qy 1090 GAGCAGAGAAGAAGCACTGTATGGTTTGAAGCTCGGAGGACGTGGAGGATGCAAGTC 1149  
Db 1176 GAGCAGAGAAGAAGCACTGTATGGTTTGAAGCTCGGAGGACGTGGAGGATGCAAGTC 1235  
  
Qy 1150 GTCTGTGGGAATTTCTGGCCATCAATGCCACAAACATGTCTGTGCTTGTGCGCGAGC 1209  
Db 1236 GTCTGTGGGAATTTCTGGCCATCAATGCCACAAACATGTCTGTGCTTGTGCGCGAGC 1295  
  
Qy 1210 CCCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTCTGACCTCATCTCATC 1269  
Db 1296 CCCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTCTGACCTCATCTCATC 1355  
  
Qy 1270 CGGAAATGCTCCAGGTTCATTTTCTGAGATTTCTCATCAGGCAACCAACACAGGAC 1329  
Db 1356 CGGAAATGCTCCAGGTTCATTTTCTGAGATTTCTCATCAGGCAACCAACACAGGAC 1415  
  
Qy 1330 CAGTTTGACTTCACTTTTGTGTAAGTTTATCGCTCAAGAAATTCAGTTTACGTGGA 1389  
Db 1416 CAGTTTGACTTCACTTTTGTGTAAGTTTATCGCTCAAGAAATTCAGTTTACGTGGA 1475  
  
Qy 1390 CACATGGAGGATGAGGACGACCTCAAGGAGGGGGAAGACGCTTTGGGCAATT 1449  
Db 1476 CACATGGAGGATGAGGACGACCTCAAGGAGGGGGAAGACGCTTTGGGCAATT 1535  
  
Qy 1450 TGCAGCAGCCACCCCTCTGCTGTGCAACCGTCTCCACAGCTCTCTGGAATCGGACGG 1509  
Db 1536 TGCAGCAGCCACCCCTCTGCTGTGCAACCGTCTCCACAGCTCTCTGGAATCGGACGG 1595  
  
Qy 1510 GAGTCTGTGACAGCCCTGCCATGAGTCAAGTTC 1546  
Db 1596 GAGTCTGTGACAGCCCTGCCATGAGTCAAGTTC 1632

RESULT 15  
AAS77731  
ID AAS77731 standard; cDNA; 1570 BP.  
XX  
AC AAS77731;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #13535.  
XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX Homo sapiens.  
XX WO200175067-A2.  
XX 11-OCT-2001.  
XX 30-MAR-2001; 2001WO-US008631.  
XX 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX (HYSE-) HYSEQ INC.  
XX Drmanac RT, Liu C, Tang YT;  
XX WPI; 2001-639362/73.  
DR P-PSDB; ABG13544.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
PS Claim 1; SEQ ID NO 13535; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 1570 BP; 349 A; 446 C; 456 G; 319 T; 0 U; 0 Other;  
  
Query Match 46.6%; Score 752.4; DB 5; Length 1570;  
Best Local Similarity 99.9%; Pred. No. 5.8e-185;  
Matches 753; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 790 GGGGACTCGCTGGCCATGGATGTCTCTCAGTCACACACACACACTCTTCGGTAC 849  
Db 695 GGGGACTCGCTGGCCATGGATGTCTCTCAGTCACACACACACACTCTTCGGTAC 754  
  
Qy 850 TCCGTGTCCCTGCTGGGTACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAA 909  
Db 755 TCCGTGTCCCTGCTGGGTACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAA 814  
  
Qy 910 CGGTGGTGGGTCTTGGCAGATACGACTTTTCAGGTTTAAAGACTTCTCTCCACAC 969  
Db 815 CGGTGGTGGGTCTTGGCAGATACGACTTTTCAGGTTTAAAGACTTCTCTCCACAC 874  
  
Qy 970 TGCTATGAGGACAGTGTCTCTCTCTCCACACACACAGGTGGATCTCCAGGGAT 1029  
Db 875 TGCTATGAGGACAGTGTCTCTCTCTCCACACACACAGGTGGATCTCCAGGGAT 934  
  
Qy 1030 AGGAAGCCCTCCGGGAGGATGCTTTGTTGACGGAAGCAAGCAGCAGCTGGAGG 1089

```

|||||
Db 935 AGGAGCCCTGCGGGCAGGATGCTTTGTTTCAGGCAAGCAGCAGCTGGAGGAG 994
QY 1090 GAGCAGAGAAAGCACTGTATGGTTTGGAAAGCTGCGGAGGACGTGGAGGAGTGGCAAGTC 1149
Db 995 GAGCAGAGAAAGCACTGTATGGTTTGGAAAGCTGCGGAGGACGTGGAGGAGTGGCAAGTC 1054
QY 1150 GTCTGTGGGAAGTTTCTGCGCCATCAATGSCCAAAACATGTCTGTGCTTGTGCGCCGAGC 1209
Db 1055 GTCTGTGGGAAGTTTCTGCGCCATCAATGSCCAAAACATGTCTGTGCTTGTGCGCCGAGC 1114
QY 1210 CCCAGGGGCTCTCCCGGCTGCCACATTGGGAGACGGGTCTTCTGACCTCATCTCATC 1269
Db 1115 CCCAGGGGCTCTCCCGGCTGCCACATTGGGAGACGGGTCTTCTGACCTCATCTCATC 1174
QY 1270 CGGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAACCAACCCAGCAGGAC 1329
Db 1175 CGGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAACCAACCCAGCAGGAC 1234
QY 1330 CAGTTTGACTTCACTTTTGTGAGTTTATCGGTCAGAAATTCAGTTTACGTGGAAG 1389
Db 1235 CAGTTTGACTTCACTTTTGTGAGTTTATCGGTCAGAAATTCAGTTTACGTGGAAG 1294
QY 1390 CACATGGAGGATGAGCAGCAGACCTCAAGGAGGGGGGAGAGCGCTTTGGGCACATT 1449
Db 1295 CACATGGAGGATGAGCAGCAGACCTCAAGGAGGGGGGAGAGCGCTTTGGGCACATT 1354
QY 1450 TGCAGCAGCCACCCCTCCTGTGCTGTCACCCGTCTCCAACAGCTCCTGGAACTGCGACGGG 1509
Db 1355 TGCAGCAGCCACCCCTCCTGTGCTGTCACCCGTCTCCAACAGCTCCTGGAACTGCGACGGG 1414
QY 1510 GAGGTCCTGACAGCCCTGCCATCGAGGTCAGAG 1543
Db 1415 GAGGTCCTGACAGCCCTGCCATCGAGGTCAGGG 1448
```

Search completed: September 5, 2005, 20:25:17  
Job time : 939.933 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2005, 09:12:32 ; Search time 296.284 Seconds  
(without alignments)  
8913.583 Million cell updates/sec

Title: US-10-631-958-9  
Perfect score: 1614  
Sequence: 1 atg999ggcagcggggggcgc.....agccagactcacacagctga 1614

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1612.4	99.9	4432	4	US-09-774-528-148
2	100.6	6.2	1084	4	US-09-270-767-15155
3	99	6.1	2064	4	US-09-270-767-14306
4	69.8	4.3	498	4	US-09-893-737-59
5	67.8	4.2	1857	4	US-09-970-516-3
6	67.8	4.2	2380	4	US-09-817-676A-13
7	65.6	4.1	901	4	US-09-270-767-30448
8	61.2	3.8	2698	4	US-09-817-676A-11
9	51.2	3.2	1155	4	US-09-970-516-1
10	51.2	3.2	1205	4	US-09-959-897-1
11	51.2	3.2	1783	4	US-09-949-016-1155
12	49.8	3.1	666	2	US-08-875-034A-1
13	49.8	3.1	1875	4	US-09-614-221A-399
14	49.2	3.0	1611	4	US-09-248-796A-1756
15	48.4	3.0	1149	4	US-09-970-516-5
16	44.2	2.7	1533	4	US-09-205-258-90
17	43.8	2.7	7218	1	US-08-232-463-14
18	43.4	2.7	1050	4	US-09-252-991A-1269
19	43.4	2.7	1308	4	US-09-252-991A-1179
20	43	2.7	3071	4	US-09-949-016-5812
21	43	2.7	62908	4	US-09-949-016-17554
22	42.8	2.7	36519	3	US-08-923-137-2
23	42.2	2.6	430	4	US-09-621-976-16656
24	42.2	2.6	115963	4	US-09-949-016-12298
25	41.2	2.6	1318	3	US-09-125-891-1
26	41.2	2.6	2636	3	US-09-125-891-3
27	40.8	2.5	1014	4	US-09-902-540-9500

C	28	40.8	2.5	14101	4	US-09-902-540-1080	Sequence 1080, Ap
	29	40.8	2.5	112507	4	US-09-949-016-12420	Sequence 12420, A
	30	40.8	2.5	112507	4	US-09-949-016-12794	Sequence 12794, A
	31	40.8	2.5	112508	4	US-09-949-016-16589	Sequence 16589, A
	32	40.8	2.5	112508	4	US-09-949-016-16590	Sequence 16590, A
	33	40.6	2.5	1017	4	US-09-902-540-2472	Sequence 2472, Ap
	34	40.6	2.5	9080	4	US-09-902-540-1963	Sequence 1963, Ap
	35	40.6	2.5	15209	4	US-09-902-540-1110	Sequence 1110, Ap
	36	40.4	2.5	2016	4	US-09-252-991A-7070	Sequence 7070, Ap
	37	40.4	2.5	6855	4	US-09-252-991A-6976	Sequence 6976, Ap
	38	40.2	2.5	1012	4	US-09-270-767-12677	Sequence 12677, A
	39	40.2	2.5	1386	4	US-09-252-991A-8146	Sequence 8146, Ap
	40	40.2	2.5	1620	4	US-09-252-991A-8213	Sequence 8213, Ap
	41	40.2	2.5	2279	4	US-09-949-016-4013	Sequence 4013, Ap
	42	40.2	2.5	12449	4	US-09-949-016-15755	Sequence 15755, A
	43	40	2.5	685	4	US-09-902-331B-11	Sequence 11, Appl
	44	40	2.5	670689	4	US-09-949-016-12505	Sequence 12505, A
	45	40	2.5	670690	4	US-09-949-016-14207	Sequence 14207, A

ALIGNMENTS

RESULT 1  
US-09-774-528-148  
; Sequence 148, Application US/09774528  
; Patent No. 6743619  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Zhou, Ping  
; APPLICANT: Goodrich, Ryle  
; APPLICANT: Liu, Chonghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Zhang, Jie  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 802  
; CURRENT APPLICATION NUMBER: US/09/774,528  
; NUMBER OF SEQ ID NOS: 441  
; SOFTWARE: Pt\_FL\_genes Version 2.0  
; SEQ ID NO 148  
; LENGTH: 4432  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1129)..(2817)  
US-09-774-528-148

Query Match	99.9%	Score 1612.4;	DB 4;	Length 4432;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 1613;	Conservative	0;	Mismatches	1;
Indels	0;	Gaps	0;	
QY	1	ATGCGGGCGACGGGGCGGCGGAGCCGCTGCAATCCGTGCTGTGGTGAAGCAGCAGCGC	60	
Db	1204	ATGCGGGCGACGGGGCGGCGGAGCCGCTGCAATCCGTGCTGTGGTGAAGCAGCAGCGC	1263	
QY	61	TGCGCGCTGACCTGAGCCCGCGGGGCTCTGCTGCGCTGCTGCGGAGCGCGGGGCC	120	
Db	1264	TGCGCGCTGAGCCCTGAGCCCGCGGGGCTCTGCTGCGCTGCTGCGGAGCGCGGGGCC	1323	
QY	121	GGAGCGCGCGCCCCCGCGGGGATGCCCTGCTCTGTGCTGTATCTGAGATCATCGCCGTT	180	
Db	1324	GGAGCGCGCGCCCCCGCGGGATGCCCTGCTCTGTGCTGTATCTGAGATCATCGCCGTT	1383	

Qy	181	GAGGAAA	CAGACGTTCA	CGGGAAA	CATCAAGGC	CAGTGGAAA	ATGGCAGAAA	ATGGCAAA	ATGGAAA	AAAG	240
Db	1384	GAGGAAA	CAGACGTTCA	CGGGAAA	CATCAAGGC	CAGTGGAAA	ATGGCAGAAA	ATGGCAAA	ATGGAAA	AAAG	1443
Qy	241	CTTTACGCTTTT	TACAGTTTCA	CTGTGTGT	TAAAGAGACA	CGACGGC	ACCGCTGG	AAAGTGG	CGCG	300	
Db	1444	CTTTACGCTTTT	TACAGTTTCA	CTGTGTGT	TAAAGAGACA	CGACGGC	ACCGCTGG	AAAGTGG	CGCG	1503	
Qy	301	CAGGTGACTT	TCTGGTGT	TCCAGAGGAG	CAGCTGTGT	CACTTGT	GGCTG	CAGAC	CCCTCGG	360	
Db	1504	CAGGTGACTT	TCTGGTGT	TCCAGAGGAG	CAGCTGTGT	CACTTGT	GGCTG	CAGAC	CCCTCGG	1563	
Qy	361	GAGATGCT	TGGAGAA	CTGACGT	CCAGAC	CCAAAG	CAATTT	TACTT	GGTATTT	TATCA	420
Db	1564	GAGATGCT	TGGAGAA	CTGACGT	CCAGAC	CCAAAG	CAATTT	TACTT	GGTATTT	TATCA	1623
Qy	421	GGAGGAAA	GGACAA	AGGCGG	ATATG	AAAGAA	AGTGG	GCAC	CACTGTT	TTCAC	480
Db	1624	GGAGGAAA	GGACAA	AGGCGG	ATATG	AAAGAA	AGTGG	GCAC	CACTGTT	TTCAC	1683
Qy	481	GCCTCCAT	CACCA	CTGACAT	CATCGTT	TACTG	AAACAT	CTAAT	CAGG	CCAAAGG	540
Db	1684	GCCTCCAT	CACCA	CTGACAT	CATCGTT	TACTG	AAACAT	CTAAT	CAGG	CCAAAGG	1743
Qy	541	TATGAGAT	TAA	CATAG	CAAA	ATA	CGACGG	CAAT	CGTCTGT	GTGGCGG	600
Db	1744	TATGAGAT	TAA	CATAG	CAAA	ATA	CGACGG	CAAT	CGTCTGT	GTGGCGG	1803
Qy	601	AGCGAGGT	GTC	CA	CGGTCT	GATTGG	GAGGAG	CGCAG	AGGAG	CGCGCGG	660
Db	1804	AGCGAGGT	GTC	CA	CGGTCT	GATTGG	GAGGAG	CGCAG	AGGAG	CGCGCGG	1863
Qy	661	CACCCCGGG	GTGCT	GGTCCC	CAGTAG	CCCTCCG	GATTGGA	ATCATTT	CCCGCAG	GGTCA	720
Db	1864	CACCCCGGG	GTGCT	GGTCCC	CAGTAG	CCCTCCG	GATTGGA	ATCATTT	CCCGCAG	GGTCA	1923
Qy	721	ACGGA	CTCGTGT	TTACT	CA	CGTGGG	CA	CCAG	CGACG	CAGAAA	780
Db	1924	ACGGA	CTCGTGT	TTACT	CA	CGTGGG	CA	CCAG	CGACG	CAGAAA	1983
Qy	781	ATCGTGT	TTGGG	GA	CTCGT	GGC	ATG	GTCTC	AGT	CCAC	840
Db	1984	ATCGTGT	TTGGG	GA	CTCGT	GGC	ATG	GTCTC	AGT	CCAC	2043
Qy	841	CTTCG	TACTC	CGTGT	CCCTG	CTCGG	CTACGG	CTTCTAC	GGGAG	ACATCAT	900
Db	2044	CTTCG	TACTC	CGTGT	CCCTG	CTCGG</					

Qy	1361	ATCCTCATCCGGAATGCTCAGAGTTCAATTTCTTGAGATTTCTCATCAGGCACACCAAC	1320
Db	2464	ATCCTCATCCGGAATGCTCCAGGTTCAAATTTTCTGAGATTTCTCTCATCAGGCACACCAAC	2523
Qy	1321	CAGCAGGACCAGTTTGAATCTTACATTTTGTGAAAGTTTATCGCGTCAAGAAATATCCAGTTT	1380
Db	2524	CAGCAGGACCAGTTTGAATCTTACATTTTGTGAAAGTTTATCGCGTCAAGAAATATCCAGTTT	2583
Qy	1381	ACGTCTGAAGCACATGAGAGGATGAGGACAGCGGACCTCAAAGGAGGGGGGAAGAAGCGCTTT	1440
Db	2584	ACGTCTGAAGCACATGAGAGGATGAGGACAGCGGACCTCAAAGGAGGGGGGAAGAAGCGCTTT	2643
Qy	1441	GGGCACATTTGTCAGCAGCCACCCCTCTGCTGTGTGACACCGTCTTCCAAACAGCTCTCTGGAAC	1500
Db	2644	GGGCACATTTGTCAGCAGCCACCCCTCTGCTGTGTGACACCGTCTTCCAAACAGCTCTCTGGAAC	2703
Qy	1501	TGCCACGGGGAGGTCGTGACAGCCCTCCATCGAGGTCAGAGTCCAGCTGCCAGCTCGTGT	1560
Db	2704	TGCCACGGGGAGGTCGTGACAGCCCTCCATCGAGGTCAGAGTCCAGCTGCCAGCTCGTGT	2763
Qy	1561	CGACTCTTTGTCAGGAGAAATTGAAGAAATCCGAAGCCAGACTCACAACAGCTGA	1614
Db	2764	CGACTCTTTGTCAGGAGAAATTGAAGAAATCCGAAGCCAGACTCACAACAGCTGA	2817
RESULT 2			
US-09-270-767-15155/c			
; Sequence 15155, Application US/09270767			
; Patent No. 6703491			
; GENERAL INFORMATION:			
; APPLICANT: Homburger et al.			
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster			
; FILE REFERENCE: File Reference: 7326-094			
; CURRENT APPLICATION NUMBER: US/09/270,767			
; CURRENT FILING DATE: 1999-03-17			
; NUMBER OF SEQ ID NOS: 62517			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 15155			
; LENGTH: 1084			
; TYPE: DNA			
; ORGANISM: Drosophila melanogaster			
US-09-270-767-15155			
Query Match 6.2%; Score 100.6; DB 4; Length 1084;			
Best Local Similarity 49.4%; Pred.No. 9.1e-17;			
Matches 290; Conservative 0; Mismatches 294; Indels 3; Gaps 1;			
Qy	401	TGGTATTATTCAACCCGCTTTGGAGGAAAAGCAAGGCAAGCGGATATATGAAGAAAG	460
Db	1028	TGGTCTTTATAACCCCTATGGAGTTCGAAGGGGAGCTCAGACCTATAGCGGCATG	969
Qy	461	TGGCACCACCTGTTCCACTTACCTCCATCACCACCTGACATCATCGTTTACTGAAACATGCTA	520
Db	968	TGAGACCTATTTTCCAGCTTCCGCGCTAGACGCCACGCTGATCACCACCTCAGAGGCAA	909
Qy	521	ATCAGGCCAAGGACACTCTGTATGAGATTAAACATAGACAATACGACGGCATCGTCTGTG	580
Db	908	ACCAAGTGAAGGACATACCTTCTTGAGCCATGATCTGGGAGTATACGATGCGGTTTGTGTG	849
Qy	581	TCGCGGAGATGGTATGTTTCAGCGAGTGTGTGACCGTCTGATTTGGAGGACGACAGAGGA	640
Db	848	TCGAGGGCATGGCACCGTAGCAGAGTTCATCAACGGACTGATATTCGTCAAATGCGAG	789
Qy	641	CGCGCGGGGTGACACAGAACCAACCCCGGGGTGTGTGTGTCCTCCAGGTAGCGCTCCGGAATG	700
Db	788	AGTTGGGACTGGACGAACAGCGGGCCACCATATCCAAAGACCGGCT---CTGCCAGTGG	732
Qy	701	GAATCATTTCCGCGAGGTCACAGCACTGCGTGTGTTACTCCACGTCGGCACACGACGACG	760
Db	731	GTGTGATTTCCGCTCTGGCAGCACCGACCAATTTGCGTATAGTATGACAGGCACGCGCGATG	672
Qy	761	CAGAAACCTCGGCGCTGCATATCGTTTGTGGGACTCGCTGGGCAATCGATGTGTCTCAG	820

Db 671 TGAGGACAGCGGTATCCATGTGATTCGGCCAGCATCGGGATTGGATGTGTCAGTG 612  
Qy 821 TCCACCAACAAGCACATCTCTCGTACTCCGTGTCTCGTGTCCGGCTACGGCTTCTACG 880  
Db 611 TGAGCAATGGCCAGTCCCTGCTCAGATTCTGTGCCAGTGTCTGAGCTACGGGTACCTGG 552  
Qy 881 GGGACATCATCAAGACAGCTGAGAAACGGTGGTGGTCTTGCAGATACGACTTTT 940  
Db 551 GCGATGTGCCAGCCAGACGGAGAACTACCGCTGGATGGACCGCGCGGTACGAGTACA 492  
Qy 941 CAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTG 987  
Db 491 GTGGCTCAAGGCTTCTCGAATAATCGCGGTATGAGCGCCGAAGTG 445

RESULT 3  
US-09-270-767-14306  
; Sequence 14306, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 14306  
; LENGTH: 2064  
; TYPE: DNA  
; ORGANISM: *Drosophila melanogaster*  
US-09-270-767-14306

Query Match 6.1%; Score 99; DB 4; Length 2064;  
Best Local Similarity 49.2%; Pred. No. 3.4e-16;  
Matches 289; Conservative 0; Mismatches 295; Indels 3; Gaps 1;

Qy 401 TGGTATTATCAACCGTTTGGAGAAAGGACAGGCAAGCGGATATATGAAGAAAG 460  
Db 620 TGGTCTTTATAAACCCCTATGGAGTTCGCAAGCGGGAGCTCAGACCTATGAGCCCATG 679  
Qy 461 TGGCACCACCTGTTACCTTAGCTCCATCACCACCTGACATCATCTGTTACTGAACATGCTA 520  
Db 680 TGAGACCTATTTTCCAGCTTTCGGCGGTAGACGCCAGTGATACCACTCAGAGGGCAA 739  
Qy 521 ATCAGGCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGACGCGCATCGTCTGTG 580  
Db 740 ACCAAGTGAAGGACATCTTCTGAGCCATGATCTGGAGTATACGATGCGGTTTCTGTG 799  
Qy 581 TCGCGGAGATGGTATGTTTCAGGAGTCTGACGGTCTGATTTGGGAGGACGCAAGGA 640  
Db 800 TCGGAGCGGATGGCAGCGGTAGCAGAGGTCTATCAACGGACTGATATTCGTCAAAATCGGAG 859  
Qy 641 GCGCGGGGTCCACAGACACACCCCGGGCTGTGCTGCTCCAGTAGCCCTCCGGATTG 700  
Db 860 AGTGGGACTGGACGACACCGCCGACCATATATCCAGACCGCT---CTGCCAGTGG 916  
Qy 701 GAATCATTTCCCGAGGGTCAACGGACTGCGTGTGTTACTCCACCGTGGGACACGACGAG 760  
Db 917 GTGTGATTTCCCGTGGCAGCACGACACCATTTGCGTATAGTATGACGCGCAGCGGATG 976  
Qy 761 CAGAAACCTTCGCGCTGCAATATCGTTGTTGGGAGCTCGCTGGCCATGGATGTCTCTAG 820  
Db 977 TGAGGACAGCGGCTATCCATGTGATTCGGGCGAGCATCGGGGATTTGGATGTGTCAGTG 1036  
Qy 821 TCCACCAACAAGCACATCTCTCGTACTCGTGTCTCGTGTCCGGCTACGGCTTCTACG 880  
Db 1037 TGAGCAATGGCAGTCCCTGCTCAGATTCTGTGCCAGTGTCTGAGCTACCGGGTACCTGG 1096  
Qy 881 GGGACATCATCAAGGACAGTGAAGAAACGGTGGTGGTCTTGCAGATACGACTTTT 940  
Db 1097 GCGATGTGCGAGCCGAGCGAGAACTACCGCTGGATGGACCGCGCGGTACGAGTACA 1156

Qy 941 CAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAGGACAGTG 987  
Db 1157 GTGGCGTCAAGGCTTCTCTGAATAATCGCGCTATGAGCGCGAACTG 1203

RESULT 4  
US-09-893-737-59/c  
; Sequence 59, Application US/09893737  
; Patent No. 6822082  
; GENERAL INFORMATION:  
; APPLICANT: Sheppard, Paul O.  
; APPLICANT: Presnell, Scott R.  
; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS  
; FILE REFERENCE: 00-41  
; CURRENT APPLICATION NUMBER: US/09/893,737  
; CURRENT FILING DATE: 2001-06-28  
; PRIOR APPLICATION NUMBER: US 60/215,446  
; PRIOR FILING DATE: 2000-06-30  
; NUMBER OF SEQ ID NOS: 329  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 59  
; LENGTH: 498  
; TYPE: DNA  
; ORGANISM: *Homo sapiens*  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(498)  
US-09-893-737-59

Query Match 4.3%; Score 69.8; DB 4; Length 498;  
Best Local Similarity 91.4%; Pred. No. 1.1e-08;  
Matches 74; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 713 CAGGTCACACGAGCTGCGTGTACTCCACCGTGGGACGACGACGACGAACTCGG 772  
Db 460 CAGGTCACACGAGCTGCGTGTGACTCCACCGTGGGACGACGCGCGGCGGAAACCGCGG 401

Qy 773 CGCTGCATATCGTTGTTGGGG 793  
Db 400 CGCTGCATATCGTTGTTGGG 380

RESULT 5  
US-09-970-516-3  
; Sequence 3, Application US/09970516  
; Patent No. 6610534  
; GENERAL INFORMATION:  
; APPLICANT: No. 6610534artis AG  
; TITLE OF INVENTION: Induction of blood vessel formation through administration of  
; FILE REFERENCE: polynucleotides encoding sphingosine kinases  
; FILE REFERENCE: 4-31617  
; CURRENT APPLICATION NUMBER: US/09/970,516  
; CURRENT FILING DATE: 2001-10-04  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 1857  
; TYPE: DNA  
; ORGANISM: *Homo sapiens*  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1857)  
; OTHER INFORMATION:  
US-09-970-516-3

Query Match 4.2%; Score 67.8; DB 4; Length 1857;  
Best Local Similarity 45.4%; Pred. No. 7.7e-08;  
Matches 292; Conservative 0; Mismatches 342; Indels 9; Gaps 1;

Qy 359 GGGAGATGCTGGAGAGCTGACGTCAGACCAAGAGCATTTACTGTATTTATCAACCGGT 418  
Db 401 GGGAGATCACCCCTGACCTGCTACCTCGCGCGCGGTTGTTCTTATTTGTCATCCCT 460

```
Qy 419 TTGAGGAAAAGGACAAGGCAAGCGGATATATGAAAGAAAGTGGCAACCACTGTTACCT 478
Db 461 TTTGGGGTCGGGGCTGGCCCTGGCAGTGGTGTAAAGAACCACTGTTCCCATGATCTCTG 520
Qy 479 TAGCCTCCATCACCACCTGACATCGTTACTGTGACATGCTAAATCAGGCCAAGGACTC 538
Db 521 AAGCTGGGCTGTCTTCACTCATCCAGACAGAACGACAGAACCGCCCGGAGCTGG 580
Qy 539 TGTATGAGATTAAACATAGACAAATAACGACGCATCGTCTGTGTGGGGAGATGGTATGT 598
Db 581 TCCAGGGCTGAGCCTGAGTGAAGTGGATGCAATCGTCACGGTCTCGGGAGAGGGCTGC 640
Qy 599 TCAGCGAGTCTGTCAGCGTCTGATGAGGAGGAGACGACAGAGAGCGCCGGGTGCACAGA 658
Db 641 TCCATGAGTCTGTAACCGGGCTCCTAGATCGCCCTGACTGGGAGGAAAGCTGTGAAGATG- 699
Qy 659 ACCACCCCGGGCTGTGCTGCTCCAGTAGACCTCCGGATTGGAAATCATTTCCCGCGGGT 718
Db 700 -----CCTGTGGGCATCTCTCCCTCGGCTCGGGCAACGCGCTGGCCGGAGCAGTGA 751
Qy 719 CAAGGACTGCGTGTGTACTCCACCGTGGGCACACGACGACGAGAAACCTTCGGCGCTGC 778
Db 752 ACCAGACGGGGATTTGAGCCAGCCCTGGGCTCGACCTTGTCTCACTGCTCACTGT 811
Qy 779 ATATCGTTTGGGACTCGTGGCCATGAGTGTCTCAGTCCACCAACACAGCAC 838
Db 812 TGCTGTGCGGGGTGGTGGCCACCACTGGACCTGCTCTCGTGACGCTGGCCCTCGGGCT 871
Qy 839 TCCTTCGCTACTCGTGTCTCTGCTGGCTAGGCTTCTACGGGACATCATCAAGACA 898
Db 872 CCCGCTGTTCTCTCTGCTGTGGCTTGGGCTTGGGCTTGGCTGAGATGTGGATATCCAGA 931
Qy 899 GTGAGAAAGAAACGGTGTGGTCTTGCCAGATAGACTTTTCAGGTTTAAAGACCTTCC 958
Db 932 CGAGCGCTTCAGGGCTTGGGCTTGGGCTGCGGCTTCACTGGGACGCTGGGCTCG 991
Qy 959 TCTCCCACTGCTATGAAGGACAGTGTCTTCTCCCTCCCTGC 1001
Db 992 CCACACTGCACACTACCGGAGCGCTCTCTACCTCCCGC 1034
```

## RESULT 6

```
US-09-817-676A-13
; Sequence 13, Application US/09817676A
; Patent No. 6800470
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; FILE REFERENCE: Expression and Methods of Use Thereof
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817,676A
; PRIOR APPLICATION NUMBER: 2001-03-26
; PRIOR FILING DATE: US 60/194,318
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 2380
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1860)
; PUBLICATION INFORMATION:
; TITLE: Molecular cloning and functional characterization of a
; TITLE: novel mammalian sphingosine kinase type 2 isoform
; JOURNAL: J. Biol. Chem.
; VOLUME: 275
; ISSUE: 26
; PAGES: 19513-19520
; DATE: 2000-06-30
; DATABASE ACCESSION NUMBER: AF245447
```

```
; DATABASE ENTRY DATE: 2000-06-27
US-09-817-676A-13
Query Match 4.2%; Score 67.8; DB 4; Length 2380;
Best Local Similarity 45.4%; Pred. No. 8.8e-08;
Matches 292; Conservative 0; Mismatches 342; Indels 9; Gaps 1;
Qy 359 GGGAGATGCTGGAGAAAGCTGACCTCCAGACCAAGCAATTTACTGTGTTATTTATCAACCGT 418
Db 407 GGGAGATCACCCCTGACCTGCTACCTCGGCCGCCCGTTGCTTCTATTGGTCAATCCCT 466
Qy 419 TTGGAGAAAAGGACAAGGCAAGGGATATATGAAAGAAAGTGGCAACCACTGTTACCT 478
Db 467 TTGGGGTTCGGGGCTGGCCTGGCAGTGGTGTGAAGAACCAACGCTGTTCCCATGATCTCTG 526
Qy 479 TAGCCTCCATCACCACTGACATCATCGTTACTGAAATGCTAATCAGGCCAAGGAGACTC 538
Db 527 AAGCTGGGCTGTCTTCAACTCATCCAGACAGACGACAGAACCCCGGAGCTGG 586
Qy 539 TGTATGAGATTAAACATAGACAAATAACGACGCATCGTCTGTGTGGCGGAGATGGTATGT 598
Db 587 TCCAGGGCTGAGCCTGAGTGAAGTGGATGGCATCGTCAAGTCTCGGGAGACGGGCTGC 646
Qy 599 TCAGCGAGTGTGTCAGCGTCTGATTTGGGAGGACGACAGAGAGCGCCGGGTGCACAGA 658
Db 647 TCCATGAGTGTGCTGAACCGGGCTCCTAGATCGCCCTGACTGGGAGGAAAGCTGTGAAGATG- 705
Qy 659 ACCACCCCGGGCTGTGCTGCTCCCAAGTAGCCTCCGGATTGGAAATCATTTCCCGCAGGGT 718
Db 706 -----CCTGTGGGCATCTCTCCCTCGGCTCGGCAACGCGCTGGCGGAGCAGTGA 757
Qy 719 CAAGGACTGCGTGTGTACTCCACCGTGGGCACACGACGACGAGAAACCTTCGGCGCTGC 778
Db 758 ACCAGACGCGGGATTTGAGCCAGCCCTGGGCTCGACCTTGTGCTCAACTGCTCACTGT 817
Qy 779 ATATCGTTTGGGACTCGCTGGCCATGAGTGTCTCAGTCCACCAACACAGCAC 838
Db 818 TGCTGTGCGGGGTGGTGGCCACCCACTGACCTGCTCTCGTGACGCTGGCCCTCGGGCT 877
Qy 839 TCTTCGCTACTCGTGTCTCTGCTGGCTACGGCTTCTACGGGAGACATCATCAAGGACA 898
Db 878 CCGCTGTTCTCTTCTGCTGTGGCTTGGGCTTCTGTCAGATGTGGATATCCAGA 937
Qy 899 GTGAGAGAAACGGTGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACCTTCC 958
Db 938 GCGAGCGCTTCAGGGCTTGGGCAAGTGCCTTACACTGGGCAAGTGTGGGCTCG 997
Qy 959 TCTCCCACTGCTATGAAGGACAGTGTCTTCTCCCTCCCTGC 1001
Db 998 CCACACTGCACACTACCGGAGCGCTCTCTACCTCCCGC 1040
```

## RESULT 7

```
US-09-270-767-30448
; Sequence 30448, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30448
; LENGTH: 901
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-30448
Query Match 4.1%; Score 65.6; DB 4; Length 901;
Best Local Similarity 50.0%; Pred. No. 2.1e-07;
Matches 223; Conservative 0; Mismatches 214; Indels 9; Gaps 2;
```



Best Local Similarity 53.3%; Pred. NO. 5.5e-06;  
Matches 129; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

; Patent No. 6730480

```
; GENERAL INFORMATION:
; APPLICANT: PITSON, Stuart M
; APPLICANT: Brian, WATTENBERG W
; APPLICANT: Pu, XIA
; APPLICANT: Richard, D'ANDREA J
; APPLICANT: Jennifer, BAMBLE R
; APPLICANT: Mathew, VADAS A
; TITLE OF INVENTION: SPHINGOSINE KINASE ENZYME
; FILE REFERENCE: PITSON=1
; CURRENT APPLICATION NUMBER: US/09/959,897
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: PCT/AU00/00457
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: AU PQ 0339
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: AU PQ 1504
; PRIOR FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 1205
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (33)..(1184)
; OTHER INFORMATION:
US-09-959-897-1

Query Match          3.2%; Score 51.2; DB 4; Length 1205;
Best Local Similarity 49.6%; Pred. No. 0.0018;
Matches 131; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

Qy 398 TACTGGTATTATCAACCCGTTTGGAGGAAAGGACAAAGGCGGATATATGAAGAA 457
Db 82 TGCTGGTCTGCTGAACCCGCGCGGCGGCAAGGCAAGGCTTGACGCTCTCCGGAGTC 141
Qy 458 AAGTGGCACCACCTGTTTCACTTAGCCTCCATCACCACCTGACATCATCGTTTACTGAACATG 517
Db 142 ACGTGCAGCCCTTTTGGCTGAGGCTGAATCTCCTTCACGCTGATGCTCAGCGGC 201
Qy 518 CTAATCAGGCAAGAGACTCTGTATGAGATTACATAGACAATAACAGCGCATCGTCT 577
Db 202 GGAACACGCGGGAGCTGGTGGCGTGGAGGAGCTGGGCGCTGGGAGCGCTCTGGTGG 261
Qy 578 GTGTCGGCGGAGATGATGTTTACGCGAGGTGCTGACGCTGCTGATTGGAGGAGCGAGA 637
Db 262 TCATGCTTGGAGACGGCTGATGACGAGGTGGTGAACGGGCTCATGGAGCGGCTGACT 321
Qy 638 GGAGCGCGGGGTCGACCAGAAC 661
Db 322 GGGAGACCGCCATCCAGAAGCCCC 345

RESULT 11
US-09-949-016-1155
; Sequence 1155, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 1155
; LENGTH: 1783
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1155

Query Match          3.2%; Score 51.2; DB 4; Length 1783;
Best Local Similarity 49.6%; Pred. No. 0.0022;
Matches 131; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

Qy 398 TACTGGTATTATCAACCCGTTTGGAGGAAAGGACAAAGGCGGATATATGAAGAA 457
Db 386 TGCTGGTCTGCTGAACCCGCGCGGCGGCAAGGCTTGACGCTCTCCGGAGTC 445
Qy 458 AAGTGGCACCACCTGTTTCACTTAGCCTCCATCACCACCTGACATCATCGTTTACTGAACATG 517
Db 446 ACGTGCAGCCCTTTTGGCTGAGGCTGAATCTCCTTCACGCTGATGCTCAGCGGC 505
Qy 518 CTAATCAGGCAAGAGACTCTGTATGAGATTACATAGACAATAACAGCGCATCGTCT 577
Db 506 GGAACACGCGGGAGCTGGTGGCGTGGAGGAGCTGGGCGCTGGGAGCGCTCTGGTGG 565
Qy 578 GTGTCGGCGGAGATGATGTTTACGCGAGGTGCTGACGCTGCTGATTGGAGGAGCGAGA 637
Db 566 TCATGCTTGGAGACGGGCTGATGACGAGGTGGTGAACGGGCTCATGGAGCGGCTGACT 625
Qy 638 GGAGCGCGGGGTCGACCAGAAC 661
Db 626 GGGAGACCGCCATCCAGAAGCCCC 649

RESULT 12
US-08-875-034A-1
; Sequence 1, Application US/08875034A
; Patent No. 5968796
; GENERAL INFORMATION:
; APPLICANT: Bieseler, Barbara; Reinemer, Peter; Hain,
; APPLICANT: Rudiger; Mann, Karlheinz; Reif, Hans-Jorg; and
; APPLICANT: Thomzik, Jurgen Ernst
; TITLE OF INVENTION: DROXYRIBONUCLEIC ACID CODING FOR
; TITLE OF INVENTION: GLUTATHIONE-S-TRANSFERASE AND ITS
; TITLE OF INVENTION: USE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: Sprung Kramer Schaefer & Briscoe
; STREET: 660 White Plains Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-5144
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: System 7.5
; SOFTWARE: WordPerfect 3.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,034A
; FILING DATE: 16-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/00068
; FILING DATE: 10-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 01 840.0
; FILING DATE: 23-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kurt G. Briscoe
; REGISTRATION NUMBER: 33,141
; REFERENCE/DOCKET NUMBER: Bayer 9895-KGB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 332-1700
; TELEFAX: (914) 332-1844
; INFORMATION FOR SEQ ID NO: 1;
```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 666 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..663
US-08-875-034A-1

Query Match      3.1%; Score 49.8; DB 2; Length 666;
Best Local Similarity 53.3%; Pred. NO. 0.0031;
Matches 105; Conservative 0; Mismatches 99; Indels 0; Gaps 0

Qy      7  GCGACGGGGGGCGGAGCCCGCTGCAATCGCTGCTGTGGGTGAGACAGACAGCGGTGCGCC 66
        |||||
Db      265 GCGACGGCGTCCGGGGCGAAGCTGGAGGTGTGGGTGAGGTGCGACCACTTCAC 324
        |||||

Qy      67 GTGAGCCTGGAGCCCGCGCGGCTCTGCTGCGCTGCTGGCGAGCCCGGGGCGCGAGCC 126
        |||||
Db      325 CCGAACCGTCCGCGTGGTTCAGCTGCTGCTGAGCGCGCTCTCTGGGCGGGCGCC 384
        |||||

Qy      127 GGGCCCCCGCGCGGGATGCGCTGCTCTGTGCTGTATCTGAGATCATCGCGCTTGAGGAA 186
        |||||
Db      385 GACCGCGGGTGGTGGAGAACACCGCGAGCAGCTCGGCCAAGGTGCTCGAGCTGTACGAG 444
        |||||

Qy      187 ACAGACGTTTACGGGAA 203
        |||||
Db      445 GCGACACTGGCCCGCAA 461
        |||||

```

```

RESULT 13
US-09-614-221A-399
; Sequence 399, Application US/09614221A
; Patent No. 6723837
; GENERAL INFORMATION:
; APPLICANT: Karunanandaa, Balasulojini
; APPLICANT: Yu, Jaehyuk
; APPLICANT: Kishore, Ganesh M.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED
; TITLE OF INVENTION: WITH STEROL SYNTHESIS AND METABOLISM
; FILE REFERENCE: 16516.075
; CURRENT APPLICATION NUMBER: US/09/614.221A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/142,981
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 626
; SEQ ID NO 399
; LENGTH: 1875
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-614-221A-399

Query Match          3.1%; Score 49.6; DB 4; Length 1875;
Best Local Similarity 49.6%; Pred. No. 0.006;
Matches 127; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

Qy      398  TACTGGTATTTATCAACCGGTTTCGAGGAAAAGGACAAAGCGATATATGAAAGAA 457
Db      686  TATTAGTCATTTATTATATCCCAACGGTGGTAAAGGTACTGCTTAAAAATTTATTCCTGCAA 745
Qy      458  AAGTGGGACCACTGTTACCTTAGCTCCATCACCATGCACTCATCTGTTACTGAACATG 517
Db      746  AAGCAAGCGCCAATACTACTAGTGAAAGTGGCTGCAAAATAGAAATTTGCATACACAAATATG 805
Qy      518  CTAATCAGGCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGACGGCATCGTCT 577
Db      806  CCGCGTCAACGCCATCGATATTGCCAAAGATTTAGATATACAGCAAAATACGATACCATTTGCAT 865
Qy      578  GTGTGCGGGAGATGGTATGTTTCACGCGAGGTGCTGCACGGTCTCGATTGGGAGGACGCAGA 637

```

Db	866	GTGCTCGGCTGATGGTATTCCATACGAAGTAATTAATGGGCTTTATAGAAGACCCGACA	925
Qy	638	GGAGGCGCGGGTCTGA	653
Db	926	GAGTGGATGCGTTCAA	941

RESULT 14

US-09-248-796A-1756

; Sequence 1756, Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDID

; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 1756

; LENGTH: 1611

; TYPE: DNA

; ORGANISM: Candida albicans

US-09-248-796A-1756

	Query Match	3.0%	Score 49.2;	DB 4;	Length 1611;
	Best Local Similarity	51.4%;	Pred. No. 0.007;		
	Matches 114;	Conservative 0;	Mismatches 108;	Indels 0;	Gaps 0;
Qy	396	TTTACTGGTATTATTAACAACCGGTTTGAGAGAAAGGACAAAGCGGATATATGAAG	455		
Db	492	TATTTTGGTATTGATAAACCGGATGGGCGCAAGGACAGCCAAAACATATTACAAAA	551		
Qy	456	AAAAGTGGCAACCACTGTGTCACTTAGCCTCCATCAACCACTGCATCATCTGTACTGAACA	515		
Db	552	TAAAAATCTTACCAATATTTACAAGCGGTCGTGCTAAATGTTTACGTATTTTGAACATAATA	611		
Qy	516	TGCTAATTCAGGCCCAAGGAGACTCTGTATGAGATTTAACATAGACAAATACGACGCGCATCGT	575		
Db	612	TCATGGACACGCCCACTGAGATTCCGGCTGAGCTAGATGCTCAATGATTATGATATATTTGT	671		
Qy	576	CTGTGTCGGCGGAGATGGTATGTTTCACGCGAGGTGCTGCACGG	617		
Db	672	TTGTGTGTTCTGGCGATGGGATACCTCATGAAGTTTATCAATGG	713		

```

RESULT 15
US-09-970-516-5
; Sequence 5, Application US/09970516
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; TITLE OF INVENTION: polynucleotides encoding sphingosine kinases
; FILE REFERENCES: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1149
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1149)
; OTHER INFORMATION:
US-09-970-516-5

```

```
Query Match      3.0%; Score 48.4; DB 4; Length 1149;
Beet Local Similarity 50.0%; Pred. No. 0.0097;
Matches 121; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

Qy 398 TACTGGTATTATCAACCCGTTTGGAGGAAAGGACAAAGGCAAGCGGATATATGAAAGAA 457
    ||||| | | ||||| | | ||||| | | ||||| | | ||||| | | ||||| | |
Db 50 TGCTGGTGCTGCTGAAACCCCGAGGTGGCAAGGGCAAGGCTCTGCAGCTCTTCCAGAGCC 109

Qy 458 AAGTGGCACCACTGTTACCTTAGCCTCCATCACCACCTGACATCATCGTTACTGAACATG 517
    ||||| | | ||||| | | ||||| | | ||||| | | ||||| | | ||||| | |
Db 110 GTGTGCAGCCCTTCTTGAGGAGGAGAGATAACCTTTAAACTGATACACCGAACGGA 169

Qy 518 CTAATCAGGCCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGCGCATCGTCT 577
    ||||| | | ||||| | | ||||| | | ||||| | | ||||| | | ||||| | |
Db 170 AGACCATGCCAGGAGCTGTGTGTCAGAGGAGTTGGGTCACTGGGACGCCCTGGCAG 229

Qy 578 GTGTCGGCGGAGATGGTATGTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGCAGA 637
    ||||| | | ||||| | | ||||| | | ||||| | | ||||| | | ||||| | |
Db 230 TCATGTCGGGTGATGGTCTGATGCATGAGGTGGTGAATGGGCTAATGGAAACGCCCGACT 289

Qy 638 GG 639
    ||
Db 290 GG 291
```

Search completed: September 5, 2005, 15:12:16  
Job time : 301.284 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2005, 12:52:06 ; Search time 1173.31 Seconds  
(without alignments)  
9008.663 Million cell updates/sec

Title: US-10-631-958-9  
Perfect score: 1614  
Sequence: 1 atggggggcagggggggc.....agccagactcacagctga 1614

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7338684 seqs, 3274456166 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*
- 20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq.\*
- 21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq.\*
- 22: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*
- 24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
- 25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1614	100.0	1614	10	US-09-969-896-9
2	1614	100.0	1614	19	US-10-631-958-9
3	1614	100.0	4413	10	US-09-969-896-16
4	1614	100.0	4413	19	US-10-631-958-16
5	1614	100.0	4429	19	US-10-618-941-55
6	1612.4	99.9	4432	17	US-10-120-988-148
7	1610.8	99.8	1740	18	US-10-262-511-39

8	1607.6	99.6	4463	15	US-10-315-597A-1	Sequence 1, Appli
9	1446.4	89.6	1840	22	US-09-784-810A-5	Sequence 5, Appli
10	1446.4	89.6	1840	22	US-10-876-281-5	Sequence 5, Appli
11	903	55.9	979	10	US-09-969-896-1	Sequence 1, Appli
12	903	55.9	979	10	US-10-631-958-1	Sequence 1, Appli
13	569	35.3	817	19	US-10-115-635-247	Sequence 247, App
14	475	29.4	550	10	US-09-969-896-8	Sequence 8, Appli
15	475	29.4	550	10	US-10-631-958-8	Sequence 8, Appli
16	464.4	28.8	474	10	US-09-969-896-4	Sequence 4, Appli
17	464.4	28.8	474	10	US-10-631-958-4	Sequence 4, Appli
18	320.2	19.8	329	10	US-09-969-896-5	Sequence 5, Appli
19	320.2	19.8	329	10	US-10-631-958-5	Sequence 5, Appli
20	312	19.3	522	9	US-09-784-810A-7	Sequence 7, Appli
21	312	19.3	522	22	US-10-876-281-7	Sequence 7, Appli
22	243.6	15.1	564	22	US-10-477-445-46	Sequence 46, Appli
23	205.4	12.7	339	9	US-09-783-590-5271	Sequence 5271, Ap
24	167	10.3	167	10	US-09-969-896-6	Sequence 6, Appli
25	167	10.3	167	10	US-10-631-958-6	Sequence 6, Appli
26	153	9.5	153	10	US-09-969-896-7	Sequence 7, Appli
27	153	9.5	153	10	US-10-631-958-7	Sequence 7, Appli
28	150	9.3	382	9	US-09-784-810A-9	Sequence 9, Appli
29	150	9.3	382	22	US-10-876-281-9	Sequence 9, Appli
30	102.6	6.4	839	20	US-10-363-345A-21227	Sequence 21227, A
31	102.6	6.4	839	20	US-10-363-345A-21228	Sequence 21228, A
32	102.6	6.4	839	21	US-10-363-483A-21227	Sequence 21227, A
33	102.6	6.4	839	21	US-10-363-483A-21228	Sequence 21228, A
34	71.4	4.4	839	20	US-10-363-345A-21225	Sequence 21225, A
35	71.4	4.4	839	20	US-10-363-345A-21226	Sequence 21226, A
36	71.4	4.4	839	21	US-10-363-483A-21225	Sequence 21225, A
37	71.4	4.4	839	21	US-10-363-483A-21226	Sequence 21226, A
38	65.8	4.3	498	9	US-09-893-737-59	Sequence 59, Appli
39	67.8	4.2	1857	9	US-09-970-516-3	Sequence 3, Appli
40	67.8	4.2	1857	18	US-10-619-344-3	Sequence 13, Appli
41	67.8	4.2	2380	9	US-09-817-676A-13	Sequence 77, Appli
42	67.8	4.2	2380	16	US-10-354-358-77	Sequence 515, App
43	67.8	4.2	2380	19	US-10-283-975A-515	Sequence 13, Appli
44	67.8	4.2	2380	20	US-10-830-677-13	Sequence 139, App
45	67.8	4.2	2380	20	US-10-737-450-139	

ALIGNMENTS

RESULT 1  
US-09-969-896-9  
; Sequence 9, Application US/09969896  
; Publication No. US20030125533A1  
; GENERAL INFORMATION:  
; APPLICANT: Kossida, Sophia  
; TITLE OF INVENTION: Regulation of human Sphingosine  
; FILE REFERENCE: 004974.00594  
; CURRENT APPLICATION NUMBER: US/09/969,896  
; CURRENT FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: US 60/238,005  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: US 60/314,113  
; PRIOR FILING DATE: 2001-08-23  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 1614  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-969-896-9

Query Match 100.0%; Score 1614; DB 10; Length 1614;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1614; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGGGGGCGACGGGGCGGCGCGCTGCAATCCGCTGCTGGTGAAGCAGCAGCGC 60  
Db 1 ATGGGGGCGACGGGGCGGCGCGCTGCAATCCGCTGCTGGTGAAGCAGCAGCGC 60



```
QY 181 GAGGAAACAGAGCTTACCGGAAACATCAAGCAGTGGAAATGCGAGAAATGGAAG 240
DB 181 GAGGAAACAGAGCTTACCGGAAACATCAAGCAGTGGAAATGCGAGAAATGGAAG 240
QY 241 CCTTACGCTTTTACAGTTCACTGTGTAAAGAGAGACGACGCAACGCTGGAAGTGGCG 300
DB 241 CCTTACGCTTTTACAGTTCACTGTGTAAAGAGAGACGACGCAACGCTGGAAGTGGCG 300
QY 301 CAGGTGAATTTCTGTGTCCAGAGGACGCTGTGTCACTGTGTGCTGCAGACCCCTGCGG 360
DB 301 CAGGTGAATTTCTGTGTCCAGAGGACGCTGTGTCACTGTGTGCTGCAGACCCCTGCGG 360
QY 361 GAGATGCTGGAGAGCTGAGCTCCAGACCAAGACATTTACTGGTATTTATCAACCGTTT 420
DB 361 GAGATGCTGGAGAGCTGAGCTCCAGACCAAGACATTTACTGGTATTTATCAACCGTTT 420
QY 421 GGAGAAAGAGCAAGGCAAGCGGATATATGAAGAAAGTGGCACCACTGTTTCACTTAA 480
DB 421 GGAGAAAGAGCAAGGCAAGCGGATATATGAAGAAAGTGGCACCACTGTTTCACTTAA 480
QY 481 GCCTCCATCACCCTGACATCATCTGTTACTGAACATGCTTAATCAGGCCAAGGAGACTCTG 540
DB 481 GCCTCCATCACCCTGACATCATCTGTTACTGAACATGCTTAATCAGGCCAAGGAGACTCTG 540
QY 541 TATGAGATTAACATAGACAAATACGACGGCATCGTCTGTCTCGGCGGAGATGTTATGTT 600
DB 541 TATGAGATTAACATAGACAAATACGACGGCATCGTCTGTCTCGGCGGAGATGTTATGTT 600
QY 601 AGCGAGGTGCTCACGGTCTGATCGGAGGACGACAGAGCGCGCGGGTCCAGCAGAAC 660
DB 601 AGCGAGGTGCTCACGGTCTGATCGGAGGACGACAGAGCGCGGGTCCAGCAGAAC 660
QY 661 CACCCCCGGGCTGTGTGTCCTCCAGTAGTCCGATTCGGATTTGAATCAITCCCGCAGGGTCA 720
DB 661 CACCCCCGGGCTGTGTGTCCTCCAGTAGTCCGATTCGGATTTGAATCAITCCCGCAGGGTCA 720
QY 721 ACAGGACTGGTGTGTACTTCAACGTTGGGACGACAGGACGACAGAACCTCGGCGTGCAT 780
DB 721 ACAGGACTGGTGTGTACTTCAACGTTGGGACGACAGGACGACAGAACCTCGGCGTGCAT 780
QY 781 ATCGTGTGTTGGGACTCGCTGCGCCATGGATGTCTCAGTCCACCAACAGCACATC 840
DB 781 ATCGTGTGTTGGGACTCGCTGCGCCATGGATGTCTCAGTCCACCAACAGCACATC 840
QY 841 CTTGCTACTCCTGTGCTCGGCTACCGCTTCTACGGGACATCATCAAGGACAGT 900
DB 841 CTTGCTACTCCTGTGCTCGGCTACCGCTTCTACGGGACATCATCAAGGACAGT 900
QY 901 GAGAAAGAACGGTGTGGTCTTGCAGATACGACTTTTCAAGTTTAAAGACCTTCCTC 960
DB 901 GAGAAAGAACGGTGTGGTCTTGCAGATACGACTTTTCAAGTTTAAAGACCTTCCTC 960
QY 961 TCCCAACCACTGTATGAAGGACAGTGTCTTCTCCCTGCAACACACAGTGGGATCT 1020
DB 961 TCCCAACCACTGTATGAAGGACAGTGTCTTCTCCCTGCAACACACAGTGGGATCT 1020
QY 1021 CCAAGGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTGAGGCAAGCAAGCAGCAG 1080
DB 1021 CCAAGGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTGAGGCAAGCAAGCAGCAG 1080
QY 1081 CTGAGGAGGAGCAGAAAGACATGTATGTTTGAAGCTGCGGAGGACGTGAGGAG 1140
DB 1081 CTGAGGAGGAGCAGAAAGACATGTATGTTTGAAGCTGCGGAGGACGTGAGGAG 1140
QY 1141 TGGCAAGTCTGTGTGGAGTTTCTGGCCATCAATGCCAACAATGCTCTGTCTTGT 1200
DB 1141 TGGCAAGTCTGTGTGGAGTTTCTGGCCATCAATGCCAACAATGCTCTGTCTTGT 1200
QY 1201 CCGCCGAGCCCGCAGGGGCTCTCCCGGCTGCCACTTGGGAGAGCGGTCTTCTCACTC 1260
DB 1201 CCGCCGAGCCCGCAGGGGCTCTCCCGGCTGCCACTTGGGAGAGCGGTCTTCTCACTC 1260
```

```
QY 1261 ATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAAC 1320
DB 1261 ATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAAC 1320
QY 1321 CAGCAGGACCAAGTTTGAATTTTCTGAAAGTTTATCGCGTCAAGAAATTCAGTTT 1380
DB 1321 CAGCAGGACCAAGTTTGAATTTTCTGAAAGTTTATCGCGTCAAGAAATTCAGTTT 1380
QY 1381 ACCTGAAAGCACATGGAGGATGAGGACAGCGACCTCAAGGAGGGGGGAAAGCGCTTT 1440
DB 1381 ACCTGAAAGCACATGGAGGATGAGGACAGCGACCTCAAGGAGGGGGGAAAGCGCTTT 1440
QY 1441 GGGCACATTTCCAGCAGCCACCCCTCTGCTGTCACCGTCTCCACAGCTCTCTGGAAC 1500
DB 1441 GGGCACATTTCCAGCAGCCACCCCTCTGCTGTCACCGTCTCCACAGCTCTCTGGAAC 1500
QY 1501 TGCCACGGGAGGCTCTGCACAGCCCTGCCATCGAGGTCCAGTCCACTGCCAGCTGTT 1560
DB 1501 TGCCACGGGAGGCTCTGCACAGCCCTGCCATCGAGGTCCAGTCCACTGCCAGCTGTT 1560
QY 1561 CGACTCTTTGACAGGAAATTTGAAGAGAAATCCGAAGCCAGACTCACAGCTGA 1614
DB 1561 CGACTCTTTGACAGGAAATTTGAAGAGAAATCCGAAGCCAGACTCACAGCTGA 1614

RESULT 3
US-09-969-896-16
; Sequence 16, Application US/09969896
; Publication No. US20030125533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: Kinase-Like Protein
; CURRENT APPLICATION NUMBER: US/09/969,896
; PRIORITY FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 4413
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-16

Query Match 100.0%; Score 1614; DB 10; Length 4413;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1614; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGGCGACGGGGGGCGGAGCCGCTGCAATCCGTTGTTGGTCAAGCAGCAGCGC 60
DB 76 ATGGGGGGCGACGGGGGGCGGAGCCGCTGCAATCCGTTGTTGGTCAAGCAGCAGCGC 135
QY 61 TGGCGGTGAGCCCTGGAGCCCGCGGGCTCTGCTGCGCTGTTGGCGAGCCCGGGGCC 120
DB 136 TGGCGGTGAGCCCTGGAGCCCGCGGGCTCTGCTGCGCTGTTGGCGAGCCCGGGGCC 195
QY 121 GGAGCGGCGCCCGCGGGCGGATGCTGCTCTGTGCTCTGTATCTGAGATCATCGCCGTT 180
DB 196 GGAGCGGCGCCCGCGGGCGGATGCTGCTCTGTGCTCTGTATCTGAGATCATCGCCGTT 255
QY 181 GAGGAAACAGAGTTTCAGGGAAACATCAAGCAGTGGAAATGCGACAAATGGAAG 240
DB 256 GAGGAAACAGAGTTTCAGGGAAACATCAAGCAGTGGAAATGCGACAAATGGAAG 315
QY 241 CTTTACGCTTTTACAGTTTCACTGTGTAAGAGAGCAGCAGCGCACCGCTGGAAGTGGCG 300
DB 316 CTTTACGCTTTTACAGTTTCACTGTGTAAGAGAGCAGCAGCGCACCGCTGGAAGTGGCG 375
QY 301 CAGGTGACTTTTCTGGTGTCCAGAGAGCAGCTGTGTCTCTGTGCTGCGAGACCCCTGCGG 360
```



```
Db 376 ||||| CAGGTGACTTTCTGGTGTCCAGAGGAGCAGCTGTGTCACTTGTGGCTGCAGACCCCTGCGG 435
Qy 361 GAGATGCTGGAGAGCTGACGTCTCAGACCAAGCAATTTACTTGTTATTTATCAACCCGTTTT 420
Db 436 GAGATGCTGGAGAGCTGACGTCTCAGACCAAGCAATTTACTTGGTATTTATCAACCCGTTTT 495
Qy 421 GGAGGAAAGGACAAGGCAAGCGGATATATGAAGAAAGTGCGCAACAATCTTTCACCTTTA 480
Db 496 GGAGGAAAGGACAAGGCAAGCGGATATATGAAGAAAGTGCGCAACAATCTTTCACCTTTA 555
Qy 481 GCCTCCATCACCTGACATCATCGTTACTGAAATGCTTAATCAGGCAAGGAGACTGTG 540
Db 556 GCCTCCATCACCTGACATCATCGTTACTGAAATGCTTAATCAGGCAAGGAGACTGTG 615
Qy 541 TATGAGATTACATAGACAAATACGACGGCATCTGTGTCTGGCGGAGATGGTATGTTTC 600
Db 616 TATGAGATTACATAGACAAATACGACGGCATCTGTGTCTGGCGGAGATGGTATGTTTC 675
Qy 601 AGCAGAGTGTCTGACCGTCTGATTGGGAGGACGACAGAGGACGCGGGGTGACACAGAAC 660
Db 676 AGCAGAGTGTCTGACCGTCTGATTGGGAGGACGACAGAGGACGCGGGGTGACACAGAAC 735
Qy 661 CACCCCGGGGTGTCTGGTCTCCAGTAGCCCTCCGGATTGGAATCATTTCCCGCAGGGTCA 720
Db 736 CACCCCGGGGTGTCTGGTCTCCAGTAGCCCTCCGGATTGGAATCATTTCCCGCAGGGTCA 795
Qy 721 AGGACTGCTGTGTACTTACACCGTGGGACGACGAGCGACGAGAACCTCGGCGCTGCAT 780
Db 796 ACGGACTGCTGTGTACTTCAACCGTGGGACGACGAGCGACGAGAACCTCGGCGCTGCAT 855
Qy 781 ATCGTGTGTGGGACTCGCTGGGACATGATGTCTCAGTCCACCAACAGACACATC 840
Db 856 ATCGTGTGTGGGACTCGCTGGGACATGATGTCTCAGTCCACCAACAGACACATC 915
Qy 841 CTTGCTACTCCGTCTCCGTCTGGGCTA CCGCTTCTACGGGACATCATCAAGGACAGT 900
Db 916 CTTGCTACTCCGTCTCCGTCTGGGCTA CCGCTTCTACGGGACATCATCAAGGACAGT 975
Qy 901 GAGAGAACCGTGTGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGCCTTCTC 960
Db 976 GAGAGAACCGTGTGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGCCTTCTC 1035
Qy 961 TCCACCACTGTATGAAGGACAGTGTCTTCTCCCTGCAACAACAACAGGTTGGGATCT 1020
Db 1036 TCCACCACTGTATGAAGGACAGTGTCTTCTCCCTGCAACAACAACAGGTTGGGATCT 1095
Qy 1021 CCAAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTTCAGGCAAGCAAGCAGCAG 1080
Db 1096 CCAAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTTCAGGCAAGCAAGCAGCAG 1155
Qy 1081 CTGGAGGAGGACAGAAAGCACTGTATGGTTTGGAGTCTGGGAGGACGTGGAGGAG 1140
Db 1156 CTGGAGGAGGACAGAAAGCACTGTATGGTTTGGAGTCTGGGAGGACGTGGAGGAG 1215
Qy 1141 TGGCAAGTCTGTGTGGGAAGTTCTGCCATCAATGCAACAACATCTCTGTGCTTGT 1200
Db 1216 TGGCAAGTCTGTGTGGGAAGTTCTGGCCATCAATGCAACAACATCTCTGTGCTTGT 1275
Qy 1201 GCGCGAGCCCCAGGGGCTCTCCCGGCTGCCCCA CTTGGGAGCGGGTCTTCTGACCTC 1260
Db 1276 GCGCGAGCCCCAGGGGCTCTCCCGGCTGCCCCA CTTGGGAGCGGGTCTTCTGACCTC 1335
Qy 1261 ATCTCTATCCGGAATGCTCAGGTTCAATTTTCTGAGATTCTCATCAGGACACCAAC 1320
Db 1336 ATCTCTATCCGGAATGCTCAGGTTCAATTTTCTGAGATTCTCATCAGGACACCAAC 1395
Qy 1321 CAGCAGGACCAAGTTTGACTTCACTTTTGTGAAGTTTATCGGTCAGAAATTCAGTTT 1380
Db 1396 CAGCAGGACCAAGTTTGACTTCACTTTTGTGAAGTTTATCGGTCAGAAATTCAGTTT 1455
Qy 1381 ACCTCGAAGCACTGGAAGATGAGGACAGCGCACTCAAGGAGGGGGGAAAGACGCTTT 1440
|||
```

```
Db 1456 ACCTCGAAGCACATGGAGGATGAGGACAGCGA CTTCAAGAGGGGGGAAAGACGCGTTT 1515
Qy 1441 GGGCACATTTTGACAGACCCACCTCTCTGTGTGTGCAACCGCTCTCCAAACAGCTCTCGAAC 1500
Db 1516 GGGCACATTTTGACAGACCCACCTCTCTGTGTGTGCAACCGCTCTCCAAACAGCTCTCGAAC 1575
Qy 1501 TGGCAACGGGAGGTCTTGACAGACCCCTGCCATCGAGGTCAGAGTCCACTGCGAGCTGGTT 1560
Db 1576 TGGCAACGGGAGGTCTTGACAGACCCCTGCCATCGAGGTCAGAGTCCACTGCGAGCTGGTT 1635
Qy 1561 CGACTCTTTTGCACGAGGAATTGAAGAAATCCGAGCCAGACTCACACAGCTGA 1614
Db 1636 CGACTCTTTTGCACGAGGAATTGAAGAAATCCGAGCCAGACTCACACAGCTGA 1689
```

## RESULT 4

US-10-631-958-16  
; Sequence 16, Application US/10631958  
; Publication No. US20040192580A1

## GENERAL INFORMATION:

APPLICANT: Koesida, Sophia

TITLE OF INVENTION: Regulation of human Sphingosine

TITLE OF INVENTION: Kinase-Like Protein

FILE REFERENCE: 004974.00594

CURRENT APPLICATION NUMBER: US/10/631.958

CURRENT FILING DATE: 2003-08-01

PRIOR APPLICATION NUMBER: US/09/969,896

PRIOR FILING DATE: 2001-10-04

PRIOR APPLICATION NUMBER: US 60/238,005

PRIOR FILING DATE: 2000-10-06

PRIOR APPLICATION NUMBER: US 60/314,113

PRIOR FILING DATE: 2001-08-23

NUMBER OF SEQ ID NOS: 16

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 16

LENGTH: 4413

TYPE: DNA

ORGANISM: Homo sapiens

US-10-631-958-16

Query Match 100.0%; Score 1614; DB 19; Length 4413;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1614; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGGCGCAGCGGGCGGGAGCGCGTCAATCCGTGCTGTGGTGAAGCAGCAGCGC 60

Db 76 ATGGGGCGCAGCGGGCGGGAGCGCGTCAATCCGTGCTGTGGTGAAGCAGCAGCGC 135

Qy 61 TGGCCCGTGAAGCCCTGGAGCCGCGGGCTCTGCTGCGCTGTGGCGGAGCCCGGGGCC 120

Db 136 TGGCCCGTGAAGCCCTGGAGCCGCGGGCTCTGCTGCGCTGTGGCGGAGCCCGGGGCC 195

Qy 121 GGAGCCGGCCCCCGGGCGGGATGCTGCTGTGCTGTATCTGAGATCATCGCGTT 180

Db 196 GGAGCCGGCCCCCGGGCGGGATGCTGCTGTGCTGTATCTGAGATCATCGCGTT 255

Qy 181 GAGGAAACAGACGTTTCAGGGAAACATCAAGCAGTGGAAATGGCAAAATGGAAAG 240

Db 256 GAGGAAACAGACGTTTCAGGGAAACATCAAGCAGTGGAAATGGCAAAATGGAAAG 315

Qy 241 CCTTACCGTTTTTACGTTTCACTGTGTAAGAGACACGACGGCACCGCTGGAAGTGGCG 300

Db 316 CCTTACCGTTTTTACGTTTCACTGTGTAAGAGACACGACGGCACCGCTGGAAGTGGCG 375

Qy 301 CAGTGACTTTTCTGGTGTCCAGAGGAGCGTGTGTCACTTGTGGCTGCAGACCTCGCG 360

Db 376 CAGTGACTTTTCTGGTGTCCAGAGGAGCGTGTGTCACTTGTGGCTGCAGACCTCGCG 435

Qy 361 GAGATGCTGGAGAGCTGACGTCCAGACCAAGCAATTTACTGGTATTTTATCAACCCGTTT 420

Db 436 GAGATGCTGGAGAGCTGACGTCCAGACCAAGCAATTTACTGGTATTTTATCAACCCGTTT 495

Qy 421 GGAGGAAAGGACAAGGCAAGCGGATATATGAAAGAAAGTGGCACCACTGTTTACCTTA 480

Db 496 GGAGGAAAGGACAGGCGGATATATGAAAGAAAGTGGACCACTGTTACCTTA 555  
QY 481 GCCTCATCACTGACATCATCTGTTACTGAACATGCTAAATCAGGCCAAAGAGACTCTG 540  
Db 556 GCCTCCATCACCACCTGACATCATCTGTTACTGAACATGCTAAATCAGGCCAAAGGAGACTCTG 615  
QY 541 TATGAGATTACATAGACAAATACAGCGCATCTGTCGCGGAGATGGTATGTT 600  
Db 616 TATGAGATTAAATAGACAAATACAGCGCATCTGTCGCGGAGATGGTATGTT 675  
QY 601 ACCGAGGTGCTCAGCGTCTGATTTGGGAGAGCGCAGAGAGCGCGGGTTCAGCAGAAC 660  
Db 676 AGCGAGGTGCTGCACGGTCTGATTTGGGAGAGCGCAGAGAGCGCGGGTTCAGCAGAAC 735  
QY 661 CACCCCGGGCTGTGTCGTCCTCCAGTAGCTCCGATTTGGAATCATTTCCCGCAGGGTCA 720  
Db 736 CACCCCGGGCTGTGTCGTCCTCCAGTAGCTCCGATTTGGAATCATTTCCCGCAGGGTCA 795  
QY 721 ACGGACTCGTGTGTTACTCCACCGTGGGCACGAGGAGCGCAGAACCTCGCGCTGCAT 780  
Db 796 ACGGACTCGTGTGTTACTCCACCGTGGGCACGAGGAGCGCAGAACCTCGCGCTGCAT 855  
QY 781 ATCGTGTGTTGGGACTCGCTGCGCCATGATGTCTCTCAGTCCACCAACAAGACACTC 840  
Db 856 ATCGTGTGTTGGGACTCGCTGCGCCATGATGTCTCTCAGTCCACCAACAAGACACTC 915  
QY 841 TTTCGCTACTCGTGTCCCTGTGCGCTACCGCTTCTACGGGGACATCATCAAGACAGT 900  
Db 916 TTTCGCTACTCGTGTCCCTGTGCGCTACCGCTTCTACGGGGACATCATCAAGACAGT 975  
QY 901 GAGAGAAACGTTGGTCTTCCAGATAGACTTTTACGTTTAAAGACCTTCCTC 960  
Db 976 GAGAGAAACGTTGGTCTTCCAGATAGACTTTTACGTTTAAAGACCTTCCTC 1035  
QY 961 TCCACCACTGTATGAAAGGACAGTCTCTTCTCCCTGCAACAACAGCGTGGATCT 1020  
Db 1036 TCCACCACTGTATGAAAGGACAGTCTCTTCTCCCTGCAACAACAGCGTGGATCT 1095  
QY 1021 CCAAGGGATAGGAAGCCCTGCGGCGAGGATGCTTTGTTGAGCAAGCAAGCAGCAG 1080  
Db 1096 CCAAGGGATAGGAAGCCCTGCGGCGAGGATGCTTTGTTGAGCAAGCAAGCAGCAG 1155  
QY 1081 CTGGAGGAGGACAGAGAAAGCACTGTATGTTTGGAGCTGCGGAGAGCGTGGAGAG 1140  
Db 1156 CTGGAGGAGGACAGAGAAAGCACTGTATGTTTGGAGCTGCGGAGAGCGTGGAGAG 1215  
QY 1141 TGGCAAGTCTGTGTTGGGAAGTTCTTGGCACAATGCCAACAATGTCTGTCTTGT 1200  
Db 1216 TGGCAAGTCTGTGTTGGGAAGTTCTTGGCACAATGCCAACAATGTCTGTCTTGT 1275  
QY 1201 CGCGGAGGCCAGGAGGCTCTCCCGGCTGCCACTTTGGGAGAGCGGTCTTCTGACCTC 1260  
Db 1276 CGCGGAGGCCAGGAGGCTCTCCCGGCTGCCACTTTGGGAGAGCGGTCTTCTGACCTC 1335  
QY 1261 ATCTCTATCCGGAATATGCTCAGGTTCAATTTTCTGAGATTCTCATCAGGACACCAAC 1320  
Db 1336 ATCTCTATCCGGAATATGCTCAGGTTCAATTTTCTGAGATTCTCATCAGGACACCAAC 1395  
QY 1321 CAGCAGGACCACTTTGACTTCACTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTT 1380  
Db 1396 CAGCAGGACCACTTTGACTTCACTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTT 1455  
QY 1381 ACCTCGAAGCACAATGGAGATGAGGACAGCGACCTCAAGGAGGGGGGGAAGAGCGCTTT 1440  
Db 1456 ACCTCGAAGCACAATGGAGATGAGGACAGCGACCTCAAGGAGGGGGGGAAGAGCGCTTT 1515  
QY 1441 GGGGCACTTTGAGAGGACACCCCTCTGCTGCTGCAAGCGTCTCAACAGCTCTCGAAC 1500  
Db 1516 GGGGCACTTTGAGAGGACACCCCTCTGCTGCTGCAAGCGTCTCAACAGCTCTCGAAC 1575  
QY 1501 TCGCAGGGGAGGTCTTGCACAGCCCTGCCATCGAGGTCAGAGTCCACTGCCAGCTGGTT 1560

Db 1576 TGCAGCGGGAGGTCTCTGCACAGCCCTGCCATCGAGGTCAGAGTCCACTGCCAGCTGGTT 1635  
QY 1561 CGACTCTTTTGCACGAGGAATTTGAAGAGAAATCCGAAGCCAGACTCACAGCTGA 1614  
Db 1636 CGACTCTTTTGCACGAGGAATTTGAAGAGAAATCCGAAGCCAGACTCACAGCTGA 1689  
RESULT 5  
US-10-618-941-55  
; Sequence 55, Application US/10618941  
; Publication No. US20040197792A1  
; GENERAL INFORMATION:  
; APPLICANT: WHYTE, DAVID  
; APPLICANT: MANNING, GERARD  
; APPLICANT: CAENEPEEL, SEAN  
; TITLE OF INVENTION: NOVEL KINASES  
; FILE REFERENCE: 034536-0321  
; CURRENT APPLICATION NUMBER: US/10/618,941  
; CURRENT FILING DATE: 2003-07-15  
; PRIOR APPLICATION NUMBER: 60/395,632  
; PRIOR FILING DATE: 2002-07-15  
; NUMBER OF SEQ ID NOS: 143  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 55  
; LENGTH: 4429  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-618-941-55  
Query Match 100.0%; Score 1614; DB 19; Length 4429;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1614; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGCGGGCGACGCGGGCGGCGAGCCGCTGCAATCCGTGCTGTGGGTGAAGCAGACGCG 60  
Db 92 ATGCGGGCGACGCGGGCGGCGAGCCGCTGCAATCCGTGCTGTGGGTGAAGCAGACGCG 151  
QY 61 TCGCGCTGAGCCCTGGAGCCCGCGGGGCTCTGTCGCGCTGGTGGCGAGCCCGGGGCC 120  
Db 152 TCGCGCTGAGCCCTGGAGCCCGCGGGGCTCTGTCGCGCTGGTGGCGAGCCCGGGGCC 211  
QY 121 GGAGCCGCGCCCGCGGCGGATGCCCTGCTCTGTGCTGTATCTGAGATCATCCCGCTT 180  
Db 212 GGAGCCGCGCCCGCGGCGGATGCCCTGCTCTGTGCTGTATCTGAGATCATCCCGCTT 271  
QY 181 GAGGAAACAGACGTTTCACGGGAAACATCAAGGCAAGTGGAAATGCGAGAAATGGAAG 240  
Db 272 GAGGAAACAGACGTTTCACGGGAAACATCAAGGCAAGTGGAAATGCGAGAAATGGAAG 331  
QY 241 CCTTACGCTTTTACAGTTCACTGTGTAAGAGAGCAGCGCACCGCTGGAAGTGGGG 300  
Db 332 CCTTACGCTTTTACAGTTCACTGTGTAAGAGAGCAGCGCACCGCTGGAAGTGGGG 391  
QY 301 CAGGTGACTTTCTCGTGTCCAGAGAGCAGCTGTGTCTTGTGCTGCAGACCTCGCG 360  
Db 392 CAGGTGACTTTCTCGTGTCCAGAGAGCAGCTGTGTCTTGTGCTGCAGACCTCGCG 451  
QY 361 GAGATGCTGGAGAGCTGACGTCCAGACCAAGCAATTTACTGGTATTTTCAACCCGTTT 420  
Db 452 GAGATGCTGGAGAGCTGACGTCCAGACCAAGCAATTTACTGGTATTTTCAACCCGTTT 511  
QY 421 GGAGAAAGAGCAAGCGCAACCGGATATATGAAAGAAAGTGGACCACTGTTTCACTTA 480  
Db 512 GGAGAAAGAGCAAGCGCAACCGGATATATGAAAGAAAGTGGACCACTGTTTCACTTA 571  
QY 481 GCCTCCATCACCCTGACATCATCTGTTACTCAATGCTAATCAGGCCAAAGAGACTCTG 540  
Db 572 GCCTCCATCACCCTGACATCATCTGTTACTGAAATGCTAATCAGGCCAAAGAGACTCTG 631  
QY 541 TATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGTCGCGGAGATGGTATGTTTC 600  
Db 632 TATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGTCGCGGAGATGGTATGTTTC 691

```
Qy 601 AGCAGGAGTGTGACGGTCTGATTGGGAGGACGACGAGGAGCGCGGGTGCACCAAGAAC 660
Db 692 AGCAGGAGTGTGACGGTCTGATTGGGAGGACGACGAGGAGCGCGGGTGCACCAAGAAC 751
Qy 661 CACCCCGGGGTGTGCTGGTCCCGATGAGCTTCGGATTTGGAATTCATTCGCGAGGGTCA 720
Db 752 CACCCCGGGGTGTGCTGGTCCCGATGAGCTTCGGATTTGGAATTCATTCGCGAGGGTCA 811
Qy 721 ACGGACTGCGGTGTACTCCACGCTGGGACACGACGACGACGACAACTCGCGGCTGCAT 780
Db 812 ACGGACTGCGGTGTACTCCACGCTGGGACACGACGACGACGACAACTCGCGGCTGCAT 871
Qy 781 ATCGTTGTTGGGAGCTCGCTGCCCATGATGTGCTTCAGTCCACCAACAGCACACTC 840
Db 872 ATCGTTGTTGGGAGCTCGCTGCCCATGATGTGCTTCAGTCCACCAACAGCACACTC 931
Qy 841 CTTGCTACTCGGTGTCCCTGCTGGGCTTACGGGTCATCGGGGACATCATCAAGACAGT 900
Db 932 CTTGCTACTCGGTGTCCCTGCTGGGCTTACGGGTCATCATCAAGACAGT 991
Qy 901 GAGAAGAACGTTGGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGACCTTCCTC 960
Db 992 GAGAAGAACGTTGGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGACCTTCCTC 1051
Qy 961 TCCACCACTGCTATGAAGGACAGTGTCTTCCTCCCTGCACACACACGCTGGGATCT 1020
Db 1052 TCCACCACTGCTATGAAGGACAGTGTCTTCCTCCCTGCACACACACGCTGGGATCT 1111
Qy 1021 CCAAGGATAGAAAGCCCTGCGGGCAGGATGCTTTGTTTTCAGGCAAGCAAGCAGGAG 1080
Db 1112 CCAAGGATAGAAAGCCCTGCGGGCAGGATGCTTTGTTTTCAGGCAAGCAAGCAGGAG 1171
Qy 1081 CTGGAGGAGGACGAGAAAGACATGTAATGTTTGAAGCTGCGGAGAGCTGAGGAG 1140
Db 1172 CTGGAGGAGGACGAGAAAGACATGTAATGTTTGAAGCTGCGGAGAGCTGAGGAG 1231
Qy 1141 TGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGCTGCTGTGT 1200
Db 1232 TGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGCTGCTGTGT 1291
Qy 1201 CGCGGAGCCCAAGGGGCTCTCCCGGCTGCGCACTTTGGGAGAGGGGTCTTTCAGCTC 1260
Db 1292 CGCGGAGCCCAAGGGGCTCTCCCGGCTGCGCACTTTGGGAGAGGGGTCTTTCAGCTC 1351
Qy 1261 ATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGACACCAAC 1320
Db 1352 ATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGACACCAAC 1411
Qy 1321 CAGCAGGACCAAGTTGACTTCACTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTT 1380
Db 1412 CAGCAGGACCAAGTTGACTTCACTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTT 1471
Qy 1381 ACCTCGAAGACATGAGGATGAGACAGGACCTCAAGGAGGGGGGGAAGAGGCTTTT 1440
Db 1472 ACCTCGAAGACATGAGGATGAGACAGGACCTCAAGGAGGGGGGGAAGAGGCTTTT 1531
Qy 1441 GGGCACATTTGCAGCAGCACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
Db 1532 GGGCACATTTGCAGCAGCACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1591
Qy 1501 TCGCAGCGGGAGGTCTCTGCACAGCCCTGCCATCGAGGTTCAGATGCCAGCTGGTT 1560
Db 1592 TCGCAGCGGGAGGTCTCTGCACAGCCCTGCCATCGAGGTTCAGATGCCAGCTGGTT 1651
Qy 1561 CGACTCTTTGACGAGGAATTTGAAGAAATTCGAAGCCAGACTTCACAGCTGA 1614
Db 1652 CGACTCTTTGACGAGGAATTTGAAGAAATTCGAAGCCAGACTTCACAGCTGA 1705
```

```
RESULT 6
US-10-120-988-148
; Sequence 148, Application US/10120988
; Publication No. US20030219745A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Ren, Feiyan
; APPLICANT: Wang, Dunrui
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219745A1e1 Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 802CON
; CURRENT APPLICATION NUMBER: US/10/120,988
; PRIOR FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 09/774,528
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pf-fl_genes Version 2.0
; SEQ ID NO 148
; LENGTH: 4432
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1129)..(2817)
; US-10-120-988-148
```

```
Query Match 99.9%; Score 1612.4; DB 17; Length 4432;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1613; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 ATGGGGGCGACGGGGGCGGGAGCCGCTGCGAATCCGTCGTGTGGGTGAAGCAGCAGCGC 60
Db 1204 ATGGGGGCGACGGGGGCGGGAGCCGCTGCGAATCCGTCGTGTGGGTGAAGCAGCAGCGC 1263
Qy 61 TGGCCCTGAGCTTGGAGCCGCGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db 1264 TGGCCCTGAGCTTGGAGCCGCGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1323
Qy 121 GGAGCCGCGCCCCCGCGCGGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 1324 GGAGCCGCGCCCCCGCGCGGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1383
Qy 181 GAGAAACAGACGTTTCACGGGAAACATCAAGGACAGTGGAAATGGCAAAATGGAAAG 240
Db 1384 GAGAAACAGACGTTTCACGGGAAACATCAAGGACAGTGGAAATGGCAAAATGGAAAG 1443
Qy 241 CCTTACCGTTTACAGTTTCTGTAAGAGACACGACGCGACCGCTGGAAGTGGCG 300
Db 1444 CCTTACCGTTTACAGTTTCTGTAAGAGACACGACGCGACCGCTGGAAGTGGCG 1503
Qy 301 CAGGTGACTTTTCTGGTGTCCAGAGGAGCAGCTGTGTCACTTTGTGGCTGCAGACCTTGGG 360
Db 1504 CAGGTGACTTTTCTGGTGTCCAGAGGAGCAGCTGTGTCACTTTGTGGCTGCAGACCTTGGG 1563
Qy 361 GAGATGTGGAGAGCTGACGTCAGACCAAGCAATTTACTGGTATTATCAACCCGTTT 420
Db 1564 GAGATGTGGAGAGCTGACGTCAGACCAAGCAATTTACTGGTATTATCAACCCGTTT 1623
Qy 421 GGAGAAAGACAAAGCAAGCGGATATATGAAAGAAAGTGGCAACCACTGTTTCAACCTTA 480
Db 1624 GGAGAAAGACAAAGCAAGCGGATATATGAAAGAAAGTGGCAACCACTGTTTCAACCTTA 1683
Qy 481 GCCTTCCATCACTGACATCATCGTTTACTGAAACATGCTAATCAGGCCAAGGAGACTCTG 540
Db 1684 GCCTTCCATCACTGACATCATCGTTTACTGAAACATGCTAATCAGGCCAAGGAGACTCTG 1743
Qy 541 TATGAGATTAACTAGACAAATACAGCGCATCGTCTGTGTCGGGGAGATGGTATGTTTC 600
Db 1744 TATGAGATTAACTAGACAAATACAGCGCATCGTCTGTGTCGGGGAGATGGTATGTTTC 1803
Qy 601 AGCGAGGTGCTGACGGTCTGATTGGGAGGACGACGAGGAGCGCGGGTTCGACCAAGAC 660
Db 1804 AGCGAGGTGCTGACGGTCTGATTGGGAGGACGACGAGGAGCGCGGGTTCGACCAAGAC 1863
```



QY	121	GGAGCCGGCGCCCCCGGGGGGATGCGCTGCTCTGTGTGCTGTATCTGAGATCATCTGCGCGTT	180
DB	196	GGAGCCGGCGCCCCCGGTGTGTGCTGTGTCTGTGTCTGTATCTGAGATCATCTGCGCGTT	255
QY	181	GAGGAAACAGACGTTTCAACGGGMAACATCAAGGCAGTGGAAATGGCAGAAAAATGGAAAG	240
DB	256	GAGGAAACAGACGTTTCAACGGGMAACATCAAGGCAGTGGAAATGGCAGAAAAATGGAAAG	315
QY	241	CTTTACGCTTTTACAGTTTCACTGTGTAAAGAGACACGACGGCACCGCTGGAAATGGGCG	300
DB	316	CTTTACGCTTTTACAGTTTCACTGTGTAAAGAGACACGACGGCACCGCTGGAAATGGGCG	375
QY	301	CAGGTGACTTTCTGGTGTCCAGAGGACAGCTGTGTCACTTGTGTGGTGTGCAGACCTTGC	360
DB	376	CAGGTGACTTTCTGGTGTCCAGAGGACAGCTGTGTCACTTGTGTGGTGTGCAGACCTTGC	435
QY	361	GAGATGCTGGAGAAAGCTGAGCTCCAGACCAAAAGCATTTACTGTGTATTATCAACCCGTTT	420
DB	436	GAGATGCTGGAGAAAGCTGAGCTCCAGACCAAAAGCATTTACTGTGTATTATCAACCCGTTT	495
QY	421	GGAGGAAAAGSACAAGGCAAGCGGATATATGAAAGAAAAGTGGCACCACTGTTTCACTT	480
DB	496	GGAGGAAAAGSACAAGGCAAGCGGATATATGAAAGAAAAGTGGCACCACTGTTTCACTT	555
QY	481	GCCTCCATCACCCTGACATCATCTTTACTTGAACATGCTAATCAGGCCAAGGAGACTCTG	540
DB	556	GCCTCCATCACCCTGACATCATCTTTACTTGAACATGCTAATCAGGCCAAGGAGACTCTG	615
QY	541	TATCAGATTAAACATAGACAATAACGCGGCATCGTCTGTGTCCGCGGAGATGTTATGTT	600
DB	616	TATCAGATTAAACATAGACAATAACGCGGCATCGTCTGTGTCCGCGGAGATGTTATGTT	675
QY	601	AGCAGGTGTCTGACCGGTCTGATTTGGAGGACGACAGGAGCGCCGGGTTCGACCAAG	660
DB	676	AGCAGGTGTCTGACCGGTCTGATTTGGAGGACGACAGGAGCGCCGGGTTCGACCAAG	735
QY	661	CACCCCGGGCTGTGCTGTGCTCCAGTAGCCTCCGGAATTCGATCATTCGCGCAGGGTCA	720
DB	736	CACCCCGGGCTGTGCTGTGCTCCAGTAGCCTCCGGAATTCGATCATTCGCGCAGGGTCA	795
QY	721	ACGACTGCGTGTGTTACTCCACCGTGGGCACACGACGACGAGAACTCCGCGCTGCAT	780
DB	796	ACGACTGCGTGTGTTACTCCACCGTGGGCACACGACGACGAGAACTCCGCGCTGCAT	855
QY	781	ATCGTTGTTGGGACTCGCTGGCCATGGAATGTGTCTCAGTTCACCAACAGCACATC	840
DB	856	ATCGTTGTTGGGACTCGCTGGCCATGGAATGTGTCTCAGTTCACCAACAGCACATC	915
QY	841	CTTCGCTACTCCGTTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAGACAGT	900
DB	916	CTTCGCTACTCCGTTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAGACAGT	975
QY	901	GAGAGAAAACGGTGGTGGGTCTTGCAGATACGACTTTTTCAGGTTTAAAGACCTTCCCT	960
DB	976	GAGAGAAAACGGTGGTGGGTCTTGCAGATACGACTTTTTCAGGTTTAAAGACCTTCCCT	1035
QY	961	TCCCAACCATCTGATGAAGGACAGTGTCCCTTCTCCCTGCACAAACACGCTGGGATCT	1020
DB	1036	TCCCAACCATCTGATGAAGGACAGTGTCCCTTCTCCCTGCACAAACACGCTGGGATCT	1095
QY	1021	CCAGGGATAGGAAGCCCTCCGGCGAGGATCTTTCTTTTGCAGGCAACGACGACGACG	1080
DB	1096	CCAGGGATAGGAAGCCCTCCGGCGAGGATCTTTCTTTTGCAGGCAACGACGACGACG	1155
QY	1081	CTGAGGAGGACGAGAAAGCACTGTATGTTTGGAAAGCTCGCGGAGGACGCTGGAGG	1140
DB	1156	CTGAGGAGGACGAGAAAGCACTGTATGTTTGGAAAGCTCGCGGAGGACGCTGGAGG	1215
QY	1141	TGGCAAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCTGTGCTGT	1200
DB	1216	TGGCAAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCTGTGCTGT	1275

Qy	1201	CGCGGAGCCCCCAGGGGGCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTCTGACCTC	1261
Db	1276	CGCGGAGCCCCCAGGGGGCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTCTGACCTC	1335
Qy	1261	ATCTCATCTCGGAAATGCTCCAGGTTCAATTTTCTGAGATTCTCATCAGGCACACCAAC	1320
Db	1336	ATCTCATCTCGGAAATGCTCCAGGTTCAATTTTCTGAGATTCTCATCAGGCACACCAAC	1395
Qy	1321	CAGCAGGACCAGTTTGACTTTTGTGTTGAAAGTTTATCGCGTCAAGAAATTCAGTTT	1380
Db	1396	CAGCAGGACCAGTTTGACTTTTGTGTTGAAAGTTTATCGCGTCAAGAAATTCAGTTT	1455
Qy	1381	ACGTGGAAGACATGAGGATGAGGACAGGCACCTCAAGAGAGGGGGGGAAGAGCGCTTT	1440
Db	1456	ACGTGGAAGACATGAGGATGAGGACAGGCACCTCAAGAGAGGGGGGGAAGAGCGCTTT	1515
Qy	1441	GGGCACATTTGCAGCAGCACCCTCTCTGCTGTGCACCGTCTCCTCAACAGCTCCTCGAAC	1500
Db	1516	GGGCACATTTGCAGCAGCACCCTCTCTGCTGTGCACCGTCTCCTCAACAGCTCCTCGAAC	1575
Qy	1501	TGCGAGCGGGAGGTCTCTGCACAGCCCTGCCATCGAGGTCAAGTTCACCTGCCAGCTGTT	1560
Db	1576	TGCGAGCGGGAGGTCTCTGCACAGCCCTGCCATCGAGGTCAAGTTCACCTGCCAGCTGTT	1635
Qy	1561	CGACTCTTTCACAGAGAAATTGAAGAAATCCGAAGCCAGACTCACACAGTGA	1614
Db	1636	CGACTCTTTCACAGAGAAATTGAAGAAATCCGAAGCCAGACTCACACAGTGA	1689
RESULT 8			
US-10-315-597A-1			
; Sequence 1, Application US/10315597A			
; Publication No. US20030162206A1			
; GENERAL INFORMATION:			
; APPLICANT: Sugiyura, Masako			
; APPLICANT: Kono, Keita			
; APPLICANT: Kohama, Takafumi			
; TITLE OF INVENTION: Ceramide Kinase and DNA Encoding It			
; FILE REFERENCE: 02658CIP/HG			
; CURRENT APPLICATION NUMBER: US/10/315,597A			
; CURRENT FILING DATE: 2002-12-10			
; PRIOR APPLICATION NUMBER: JP 2000-178039			
; PRIOR FILING DATE: 2000-06-14			
; NUMBER OF SEQ ID NOS: 4			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 1			
; LENGTH: 4463			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: unsure			
; LOCATION: 3371			
; OTHER INFORMATION:			
US-10-315-597A-1			
Query Match			
Beet Local Similarity 99.6%; Score 1607.6; DB 16; Length 4463;			
Matches 1610; Conservative 0; Mismatches 4; Indels 0; Gaps 0;			
Qy	1	ATGGGGCGCAGCGGGGCGGCGGAGCCCTGCAATTCCTGTGTGGGTGAAGACGACGCGC	60
Db	124	ATGGGGCGCAGCGGGGCGGCGGAGCCCTGCAATTCCTGTGTGGGTGAAGACGACGCGC	183
Qy	61	TGCGCCGTGAGCTTGAGACCCGCGGGCTCTGCTGCGCTGTGGCGAGCCCGGGGCC	120
Db	184	TGCGCCGTGAGCTTGAGACCCGCGGGCTCTGCTGCGCTGTGGCGAGCCCGGGGCC	243
Qy	121	GGAGCCGGCGCCCCCGCGCGGATGCTGTCTGTGCTGTATCTGAGATCATCGCGCTT	180
Db	244	GGAGCCGGCGCCCCCGCGCGGATGCTGTCTGTGCTGTATCTGAGATCATCGCGCTT	303
Qy	181	GAGGAAAACAGAGTTTCAACGGGAAACATCAAGGCGATGGGAAATGGCGAGAAATGCGAAAG	240

Db 304 GAGGAAACAGACGTTTCAACGGGAAACATCAAGCAGTGGAAATGCGAGAAATGGAAG 363  
Qy 241 CTTTACGCTTTTACAGTTCACTGTGTAAAGAGACGACGCGGACCGCTGGAAGTGGCG 300  
Db 364 CTTTACGCTTTTACAGTTCACTGTGTAAAGAGACGACGCGGACCGCTGGAAGTGGCG 423  
Qy 301 CAGGTGACTTTCTGGTGTCCAGAGGAGCAGCTGTGTCACTTGTGGCTGCAGACCTTGGG 360  
Db 424 CAGGTGACTTTCTGGTGTCCAGAGGAGCAGCTGTGTCACTTGTGGCTGCAGACCTTGGG 483  
Qy 361 GAGATGCTGGAAGCTGACGTCAGACCAAAAGCAATTTACTGTGTATTTATCAACCGTTT 420  
Db 484 GAGATGCTGGAAGCTGACGTCAGACCAAAAGCAATTTACTGTGTATTTATCAACCGTTT 543  
Qy 421 GGAGGAAAGGACAGGCGGATATATGAAAGAAAGTGGACCACTGTTCACCTTA 480  
Db 544 GGAGGAAAGGACAGGCGGATATATGAAAGAAAGTGGACCACTGTTCACCTTA 603  
Qy 481 GCCTCCATCACCACCTGACATCATCTGTACTGAACATGCTAATCAGGCCAAGGAGACTGT 540  
Db 604 GCCTCCATCACCACCTGACATCATCTGTACTGAACATGCTAATCAGGCCAAGGAGACTGT 663  
Qy 541 TATGAGATTAACATAGACAAATACGACGGCATCGTCTGTGTGCGGCGAGATGTTATGTT 600  
Db 664 TATGAGATTAACATAGACAAATACGACGGCATCGTCTGTGTGCGGCGAGATGTTATGTT 723  
Qy 601 ACGGAGTGTCTGACGCTGATTTGGGAGGACGAGAGAGCGCGGGTGCACGAGAAC 660  
Db 724 ACGGAGTGTCTGACGCTGATTTGGGAGGACGAGAGAGCGCGGGTGCACGAGAAC 783  
Qy 661 CACCCCGGGGTGTGTGTCGCCAGTAGCTCCGCTCCGATTTGGAATCATTTCCCGCAGGTCA 720  
Db 784 CACCCCGGGGTGTGTGTCGCCAGTAGCTCCGCTCCGATTTGGAATCATTTCCCGCAGGTCA 843  
Qy 721 ACGGACTGTGTGTACTTCCACCGTGGGACGACGCGAGAGAACTTCCGCGCTGCAT 780  
Db 844 ACGGACTGTGTGTACTTCCACCGTGGGACGACGCGAGAGAACTTCCGCGCTGCAT 903  
Qy 781 ATCGTTTGGGAGCTCGTGGCCATGATGTGCTCAGTCCACCAACAGCACATC 840  
Db 904 ATCGTTTGGGAGCTCGTGGCCATGATGTGCTCAGTCCACCAACAGCACATC 963  
Qy 841 CTTCCGCTACTCGTGTCCGCTGCGGTACGCTTCTACGGGACATCATCAAGCAGT 900  
Db 964 CTTCCGCTACTCGTGTCCGCTGCGGTACGCTTCTACGGGACATCATCAAGCAGT 1023  
Qy 901 GAGAGAAACCGTGTGGTCTTGCCAGATACGACTTTTCAAGTTTAAAGACCTTCTC 960  
Db 1024 GAGAGAAACCGTGTGGTCTTGCCAGATACGACTTTTCAAGTTTAAAGACCTTCTC 1083  
Qy 961 TCCGACCACTGTATGAAGGACAGTGTCTTCTCCTGCAACACACGCTGGATCT 1020  
Db 1084 TCCGACCACTGTATGAAGGACAGTGTCTTCTCCTGCAACACACGCTGGATCT 1143  
Qy 1021 CCAAGGATAGGAAGCCCTGCGGCGAGATGCTTTGTTGAGGAGCAGTGGAGAG 1080  
Db 1144 CCAAGGATAGGAAGCCCTGCGGCGAGATGCTTTGTTGAGGAGCAGTGGAGAG 1203  
Qy 1081 CTGGAGGAGGACAGAGAAACACTGTATGTTTGGAGCTGCGGAGCAGTGGAGAG 1140  
Db 1204 CTGGAGGAGGACAGAGAAACACTGTATGTTTGGAGCTGCGGAGCAGTGGAGAG 1263  
Qy 1141 TGGCAAGTGTCTGTGGGAGTTTCTGGCCATCAATGCCAACAAATGTCTGTCTTGT 1200  
Db 1264 TGGCAAGTGTCTGTGGGAGTTTCTGGCCATCAATGCCAACAAATGTCTGTCTTGT 1323  
Qy 1201 CCGCGGAGCCGAGGGCTCTCCCGGCTGCCACTTGGGAGCGGTCTTCTGACCTC 1260  
Db 1324 CCGCGGAGCCGAGGGCTCTCCCGGCTGCCACTTGGGAGCGGTCTTCTGACCTC 1383  
Qy 1261 ATCCTCATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGACACCAAC 1320  
Db 1384 ATCCTCATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGACACCAAC 1443

Qy 1321 CAGCAGGACCAAGTTTGACTTCACTTTTGTGAAGTTTATCGCTCAAGAAATTCAGTTT 1380  
Db 1444 CAGCAGGACCAAGTTTGACTTCACTTTTGTGAAGTTTATCGCTCAAGAAATTCAGTTT 1503  
Qy 1381 ACGTCGAAGCACATCGAGGATGAGGACAGCGACCTCAAGGAGGGGGGGAAGCGCTTT 1440  
Db 1504 ACGTCGAAGCACATCGAGGATGAGGACAGCGACCTCAAGGAGGGGGGGAAGCGCTTT 1563  
Qy 1441 GGGCACATTTGACGAGCACCCCTCTGTGCTGACCGTCTCAACAGCTCTCTGGAAAC 1500  
Db 1564 GGGCACATTTGACGAGCACCCCTCTGTGCTGACCGTCTCTCAACAGCTCTCTGGAAAC 1623  
Qy 1501 TCGGACGGGAGGCTCTGACAGCCCTGCCATCGAGGTGAGAGTCCACTGCCAGCTGTT 1560  
Db 1624 TCGGATGGGAGGCTCTGACAGCCCTGCCATCGAGGTGAGAGTCCACTGCCAGCTGTT 1683  
Qy 1561 CGACTCTTTGACAGGAGAAATGAGAGAAATCGGAAGCCAGACTCAACAGCTGA 1614  
Db 1684 CGACTCTTTGACAGGAGAAATGAGAGAAATCGGAAGCCAGACTCAACAGCTGA 1737

## RESULT 9

US-09-784-810A-5

; Sequence 5, Application US/09784810A

; Patent No. US20020082203A1

; GENERAL INFORMATION:

; APPLICANT: RASTELLI, LUCA

; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING

; TITLE OF INVENTION: SAME

; FILE REFERENCE: 10716-08

; CURRENT APPLICATION NUMBER: US/09/784,810A

; CURRENT FILING DATE: 2001-02-14

; PRIOR APPLICATION NUMBER: 60/182,360

; PRIOR FILING DATE: 2000-02-14

; PRIOR APPLICATION NUMBER: 60/191,261

; PRIOR FILING DATE: 2000-03-22

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 5

; LENGTH: 1840

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-784-810A-5

Query Match 89.6%; Score 1446.4; DB 9; Length 1840;

Best Local Similarity 97.8%; Pred. No. 0;

Matches 1490; Conservative 0; Mismatches 1; Indels 33; Gaps 1;

Qy 124 GCGGCGCCCGCGGCGGATGCTGCTGTGCTGTATCTGAGATCATCGCGTTGAG 183  
Db 1 GCGGCGCCCGCGGCGGATGCTGCTGTGCTGTATCTGAGATCATCGCGTTGAG 60  
Qy 184 GAAACAGACGTTTCAAGGAAACATCAAGGAGTGGAAATGCGAGAAATGGAAGCT 243  
Db 61 GAAACAGACGTTTCAAGGAAACATCAAGGAGTGGAAATGCGAGAAATGGAAGCT 120  
Qy 244 TACGCTTTTACAGTTTCACTGTGTAAGAGAGACGACGCGCACCGCTGGAAGTGGCGGAG 303  
Db 121 TACGCTTTTACAGTTTCACTGTGTAAGAGAGACGACGCGCACCGCTGGAAGTGGCGGAG 180  
Qy 304 GTGACTTTTCTGGTGTCCAGAGGACGAGCTGTGTCTGCTGTGCTGTCAGACCTTGGGAG 363  
Db 181 GTGACTTTTCTGGTGTCCAGAGGAGCAGCTGTGTCTGCTGTGCTGTCAGACCTTGGGAG 240  
Qy 364 ATGCTGGAGAGCTGACGTCCAGACCAAGCATTTACTGGTATTTATCAACCCGTTTGA 423  
Db 241 ATGCTGGAGAGCTGACGTCCAGACCAAGCATTTACTGGTATTTATCAACCCGTTTGA 300  
Qy 424 GGAAAGGACAAAGCAGCGGATATATGAAAGAAAGTGGCAACCACTGTTACCTTAGCC 483  
Db 301 GGAAAGGACAAAGCAGCGGATATATGAAAGAAAGTGGCAACCACTGTTACCTTAGCC 360



Qy 484 TCCATCACCACCTGACATCATCG-----TTACT 510  
Db 361 TCCATCACCACCTGACATCATCGGTAAACAAATTCATATGTTAACTATGTAGAAAGTAATTA 420  
Qy 511 GAAATGCTATATCAGGCCAAGAGACCTGTGTATGAGATTAAATAGACAAATACGACGGC 570  
Db 421 GAAATGCTATATCAGGCCAAGAGACCTGTGTATGAGATTAAATAGACAAATACGACGGC 480  
Qy 571 ATCGTCTGTGCGGGAGATGGTATGTTACAGCAGAGTGCTGCACGGTCTCATTTGGGAGG 630  
Db 481 ATCGTCTGTGCGGGAGATGGTATGTTACAGCAGAGTGCTGCACGGTCTCATTTGGGAGG 540  
Qy 631 ACGCAGAGGACGCGGGGTGCACAGAAACACCCCGGGCTGTCTGTGTTCCCGAGTAC 690  
Db 541 ACGCAGAGGACGCGGGGTGCACAGAAACACCCCGGGCTGTCTGTGTTCCCGAGTAC 600  
Qy 691 CTCGGATTGGAATCATTTCCCGCAGGGTCAACGACCTGCGTGTGTTACTCCACCGTGGC 750  
Db 601 CTCGGATTGGAATCATTTCCCGCAGGGTCAACGACCTGCGTGTGTTACTCCACCGTGGC 660  
Qy 751 ACCAGCAGCAGAAACCTCGCGCTGCATATCGTTGTTGGGACTGCTGCGCATGGAT 810  
Db 661 ACCAGCAGCAGAAACCTCGCGCTGCATATCGTTGTTGGGACTGCTGCGCATGGAT 720  
Qy 811 GTGCTCTCAGTCCACCAACAGCACACTCCTTCGCTACTCCGTGCTGCTGCGCTAC 870  
Db 721 GTGCTCTCAGTCCACCAACAGCACACTCCTTCGCTACTCCGTGCTGCTGCGCTAC 780  
Qy 871 GGCTTCTACGGGACATCATCAGGACAGTGAAGAAACGGTGTGGTCTTGCACGA 930  
Db 781 GGCTTCTACGGGACATCATCAGGACAGTGAAGAAACGGTGTGGTCTTGCACGA 840  
Qy 931 TAGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCC 990  
Db 841 TAGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCC 900  
Qy 991 TTCTCTCCTGCACAACACACGCTGGGATCTCCAAAGGGATAGAAAGCCCTGCCGGCAGGA 1050  
Db 901 TTCTCTCCTGCACAACACACGCTGGGATCTCCAAAGGGATAGAAAGCCCTGCCGGCAGGA 960  
Qy 1051 TGCTTTGTTTGCAGGCAAAAGCAGCAGCTGAGGAGGAGCAGAAAGCACTGTAT 1110  
Db 961 TGCTTTGTTTGCAGGCAAAAGCAGCAGCTGAGGAGGAGCAGAAAGCACTGTAT 1020  
Qy 1111 GGTTTGGAAGCTGCGAGGACGTGAGGAGTGGCAAGTCTGTGTTGGCAAGTTCTTGCC 1170  
Db 1021 GGTTTGGAAGCTGCGAGGACGTGAGGAGTGGCAAGTCTGTGTTGGCAAGTTCTTGCC 1080  
Qy 1171 ATCAATGCCACAACATGTCTGTGCTTGTGCGCGAGCCCCAGGGGCTCTCCCCGGCT 1230  
Db 1081 ATCAATGCCACAACATGTCTGTGCTTGTGCGCGAGCCCCAGGGGCTCTCCCCGGCT 1140  
Qy 1231 GCCCATTTGGGAGACGGGTCTTTCACCTCATCTCATCCGAAATGCTCCAGGTTCAAT 1290  
Db 1141 GCCCATTTGGGAGACGGGTCTTTCACCTCATCTCATCCGAAATGCTCCAGGTTCAAT 1200  
Qy 1291 TTTCTGAGATTTCTCATCAGGCACACCAACAGCAGGACCAAGTTTGACTTTCATTTTGT 1350  
Db 1201 TTTCTGAGATTTCTCATCAGGCACACCAACAGCAGGACCAAGTTTGACTTTCATTTTGT 1260  
Qy 1351 GAAAGTTTATCCGCTCAAGAAATTCAGTTTATAGTTCGAAGCAGCATGGAGGATGAGACAGC 1410  
Db 1261 GAAAGTTTATCCGCTCAAGAAATTCAGTTTATAGTTCGAAGCAGCATGGAGGATGAGACAGC 1320  
Qy 1411 GACCTCAAGAGGGGGGAGAGAGCGCTTTGGGCAATTTGAGCAGAGCACCCTCTCTGC 1470  
Db 1321 GACCTCAAGAGGGGGGAGAGAGCGCTTTGGGCAATTTGAGCAGAGCACCCTCTCTGC 1380  
Qy 1471 TGCTGACCGCTTCCAAACAGCTTCTGGAACTGCGACGGGAGGTCTTGCACAGCCCTGCC 1530  
Db 1381 TGCTGACCGCTTCCAAACAGCTTCTGGAACTGCGACGGGAGGTCTTGCACAGCCCTGCC 1440  
Qy 1531 ATCGAGGTCAAGAGTCCAATGCGCAGCTGGTTCGACTTTTGGACGAGGAAATTGAAGAGAAT 1590

Db 1441 ATCAGGTCAGGGTCCACTGCCAGCTGTTGCACTCTTTGCACGAGGAATTGAAGAGAAT 1500  
Qy 1591 CCGAAGCCAGACTCCACACAGCTGA 1614  
Db 1501 CCGAAGCCAGACTCCACACAGCTGA 1524  
RESULT 10  
US-10-876-281-5  
; Sequence 5, Application US/10876281  
; Publication No. US20050123942A1  
; GENERAL INFORMATION:  
; APPLICANT: RASTELLI, LUCA  
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING  
; TITLE OF INVENTION: SAME  
; FILE REFERENCE: 10716-08  
; CURRENT APPLICATION NUMBER: US/10/876,281  
; CURRENT FILING DATE: 2004-06-24  
; PRIOR APPLICATION NUMBER: US/09/784,810  
; PRIOR FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: 60/182,360  
; PRIOR FILING DATE: 2000-02-14  
; PRIOR APPLICATION NUMBER: 60/191,261  
; PRIOR FILING DATE: 2000-03-22  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 1840  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-876-281-5  
Query Match 89,6%; Score 1446.4; DB 22; Length 1840;  
Best Local Similarity 97.8%; Pred. No. 0;  
Matches 1490; Conservative 0; Mismatches 1; Indels 33; Gaps 1;  
Qy 124 GCCGCGCCCCCGCGCGGATGCCCTGCTGTGCTGTATCTGAGATCATCGCGTTGAG 183  
Db 1 GCCGCGCCCCCGCGCGGATGCCCTGCTGTGCTGTATCTGAGATCATCGCGTTGAG 60  
Qy 184 GAAACAGACGTTTACGGGAAACATCAAGGCAGTGGAAAAATGGCAAAATGGAAAGCCT 243  
Db 61 GAAACAGACGTTTACGGGAAACATCAAGGCAGTGGAAAAATGGCAAAATGGAAAGCCT 120  
Qy 244 TACGCTTTTACAGTTTCACTGTGTAAGAGACGACGACCGCTGGAAGTGGCGCAG 303  
Db 121 TACGCTTTTACAGTTTCACTGTGTAAGAGACGACGACCGCTGGAAGTGGCGCAG 180  
Qy 304 GTGACTTTCTGGTGTCCAGAGGACGCTGTGTCACTTGTGGCTGCAGACCTCGCGGAG 363  
Db 181 GTGACTTTCTGGTGTCCAGAGGACGCTGTGTCACTTGTGGCTGCAGACCTCGCGGAG 240  
Qy 364 ATGCTGGAAGAGCTGAGTCCAGACCAAGCAATTTACTGGTATTATCAACCCGTTTGA 423  
Db 241 ATGCTGGAAGAGCTGAGTCCAGACCAAGCAATTTACTGGTATTATCAACCCGTTTGA 300  
Qy 424 GGAAGAGCAAGCAAGCGGATATATGAAGAAAAAGTGGCACCACCTGTTCACTTAGCC 483  
Db 301 GGAAGAGCAAGCAAGCGGATATATGAAGAAAAAGTGGCACCACCTGTTCACTTAGCC 360  
Qy 484 TCCATCACCACCTGACATCATCG-----TTACT 510  
Db 361 TCCATCACCACCTGACATCATCGGTAAACAAATTCATATGTTAACTATGTAGAAAGTAATTA 420  
Qy 511 GAAATGCTATATCAGGCCAAGAGACCTGTGTATGAGATTAAATAGACAAATACGACGGC 570  
Db 421 GAAATGCTATATCAGGCCAAGAGACCTGTGTATGAGATTAAATAGACAAATACGACGGC 480  
Qy 571 ATCGTCTGTGCGGGAGATGGTATGTTACAGCAGAGTGCTGCACGGTCTCATTTGGGAGG 630  
Db 481 ATCGTCTGTGCGGGAGATGGTATGTTACAGCAGAGTGCTGCACGGTCTCATTTGGGAGG 540



```
QY 631 ACGCAGAGAGCGCGGGTGCACAGAACACCCCGGGCTGTGTGTCCTCCAGTAGC 690
Db 541 ACGCAGAGAGCGCGGGTGCACAGAACACCCCGGGCTGTGTGTCCTCCAGTAGC 600
QY 691 CTCGGATTGGAATCATTTCCCGCAGGGTCAACGAGTGTGCTGTACTTCCACCGTGGGC 750
Db 601 CTCGGATTGGAATCATTTCCCGCAGGGTCAACGAGTGTGCTGTACTTCCACCGTGGGC 660
QY 751 ACCAGGACGCGAAGAACCTCGGGCTGCATATCGTTTGGGGACTCGTGGCCATGTGAT 810
Db 661 ACCAGGACGCGAAGAACCTCGGGCTGCATATCGTTTGGGGACTCGTGGCCATGTGAT 720
QY 811 GTGTCTCTAGTCCACACAGCAGCAGTCTTCTGCTACTCTCGTGTCCCTGTGGGCTAC 870
Db 721 GTGTCTCTAGTCCACACAGCAGCAGTCTTCTGCTACTCTCGTGTCCCTGTGGGCTAC 780
QY 871 GGCCTTCTACGGGACATCATCAAGGACAGTGAAGAAACGCTGGTGTGGGCTTCCACAGA 930
Db 781 GGCCTTCTACGGGACATCATCAAGGACAGTGAAGAAACGCTGGTGTGGGCTTCCACAGA 840
QY 931 TAGGACTTTTCAAGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGGACAGTGTCC 990
Db 841 TAGGACTTTTCAAGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGGACAGTGTCC 900
QY 991 TTCTCTCTGCAACAACACGCTGGATCTTCCAAAGGATAGGAAGCCCTGCCGGGACAGA 1050
Db 901 TTCTCTCTGCAACAACACGCTGGATCTTCCAAAGGATAGGAAGCCCTGCCGGGACAGA 960
QY 1051 TGCTTTGTTTGCAGGCAAGCAGCAGCTGGAGGAGGACAGAAAGCACTGTAT 1110
Db 961 TGCTTTGTTTGCAGGCAAGCAGCAGCTGGAGGAGGACAGAAAGCACTGTAT 1020
QY 1111 GGTTTGGAAGCTGCGGAGAGCTGAGAGTGGAAGTGGTCTGTGGGAAGTTTCTGGCC 1170
Db 1021 GGTTTGGAAGCTGCGGAGAGCTGAGAGTGGAAGTGGTCTGTGGGAAGTTTCTGGCC 1080
QY 1171 ATCAATGCCAACAATGCTGCTGTGTCGCGGAGCCCGAGGGGCTCTCCCGGGCT 1230
Db 1081 ATCAATGCCAACAATGCTGCTGTGTCGCGGAGCCCGAGGGGCTCTCCCGGGCT 1140
QY 1231 GCCCACTTGGGAGAGCGGGTCTTCTGACCTCATCTCCGAAATGCTCCAGGTTCAAT 1290
Db 1141 GCCCACTTGGGAGAGCGGGTCTTCTGACCTCATCTCCGAAATGCTCCAGGTTCAAT 1200
QY 1291 TTTCTGAGATTCTCATAGGCAACCAACGAGAGGACCAAGTTTGAATTCTTTTGT 1350
Db 1201 TTTCTGAGATTCTCATAGGCAACCAACGAGAGGACCAAGTTTGAATTCTTTTGT 1260
QY 1351 GAAGTTTATCGGTCAGAAATTCAGTTTACGTCGAAGCACAACGAGGATGAGCAGC 1410
Db 1261 GAAGTTTATCGGTCAGAAATTCAGTTTACGTCGAAGCACAACGAGGATGAGCAGC 1320
QY 1411 GACCTCAAGAGGGGGGGAAGAGCGCTTTTGGGCAATTTGAGCAGACCAACCTCTGCG 1470
Db 1321 GACCTCAAGAGGGGGGGAAGAGCGCTTTTGGGCAATTTGAGCAGACCAACCTCTGCG 1380
QY 1471 TGCTGACCGTCTCAACAGCTCTTGAATCTGGAACGAGGAGGTTCTTGACAGCCCTGCC 1530
Db 1381 TGCTGACCGTCTCAACAGCTCTTGAATCTGGAACGAGGAGGTTCTTGACAGCCCTGCC 1440
QY 1531 ATCGAGGTCAGAGTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1590
Db 1441 ATCGAGGTCAGAGTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
QY 1591 CCGAAGCCAGACTCACACAGCTGA 1614
Db 1501 CCGAAGCCAGACTCACACAGCTGA 1524
```

RESULT 11

US-09-969-896-1

; Sequence 1, Application US/09969896

; Publication No. US20030125533A1

```
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: Kinase-Like Protein
; CURRENT APPLICATION NUMBER: US/09/969,896
; PRIORITY FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 979
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-1
```

Query Match 55.9%; Score 903; DB 10; Length 979;

Best Local Similarity 96.6%; Pred. No. 1.7e-268;

Matches 946; Conservative 0; Mismatches 0; Indels 33; Gaps 1;

```
QY 387 ACCAAGCATTTACTGTATTATCAACCCGTTTGGAGAAAGCAAGGCAAGCGGAT 446
Db 1 ACCAAGCATTTACTGTATTATCAACCCGTTTGGAGAAAGCAAGGCAAGCGGAT 60
QY 447 ATATGAAAGAAAGTGGCACCACCTGTTCACTTACCTTCCATCACCCTGACATCATCG- 505
Db 61 ATATGAAAGAAAGTGGCACCACCTGTTCACTTACCTTCCATCACCCTGACATCATCG 120
QY 506 -----TTACTGAAACATGCTATCAGGCCAAGA 533
Db 121 TAACAAATTCATGTTAACTATGTAGAAAGTAACTGAAACATGCTAACTAGGCCAAGA 180
QY 534 GACTCTGTATCAGATTAAACATAGACAAATACAGGGCATCGTCTGTGTCGGCGGAGATGG 593
Db 191 GACTCTGTATGAGATTAAACATAGACAAATACAGGGCATCGTCTGTGTCGGCGGAGATGG 240
QY 594 TATGTTACAGGAGTGTCTGACGGTCTGATTGGGAGGACGACAGAGGAGCGCCGGGGTGA 653
Db 241 TATGTTACAGGAGTGTCTGACGGTCTGATTGGGAGGACGACAGAGGAGCGCCGGGGTGA 300
QY 654 CCAGAACCAACCCCGGGCTGTGTCGCCAGTAGACCTCCGGATTGGAATCATTTCCGCG 713
Db 301 CCAGAACCAACCCCGGGCTGTGTCGCCAGTAGACCTCCGGATTGGAATCATTTCCGCG 360
QY 714 AGGGTCAACGAGTGTGTTACTCCACCGTGGGACACGACGACGAGAAACCTTCGCG 773
Db 361 AGGGTCAACGAGTGTGTTACTCCACCGTGGGACACGACGACGAGAAACCTTCGCG 420
QY 774 GCTGCATATCGTTGTTGGGAGTCTCGTGGCCATGATGTCTCTCAGTCCACCAACAG 833
Db 421 GCTGCATATCGTTGTTGGGAGTCTCGTGGCCATGATGTCTCTCAGTCCACCAACAG 480
QY 834 CACACTCTTCGCTACTCCGTTGTCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAA 893
Db 481 CACACTCTTCGCTACTCCGTTGTCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAA 540
QY 894 GGACAGTCAGAGAAACCGTGTGGGTCTTGGCAGATACGACTTTTTCAGGTTTAAAGAC 953
Db 541 GGACAGTCAGAGAAACCGTGTGGGTCTTGGCAGATACGACTTTTTCAGGTTTAAAGAC 600
QY 954 CTTCTCTCCCAACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCAACACACCGGT 1013
Db 601 CTTCTCTCCCAACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCAACACACCGGT 660
QY 1014 GGGATCTCCAAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTGTCAGGCAAAAGCAA 1073
Db 661 GGGATCTCCAAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTGTCAGGCAAAAGCAA 720
QY 1074 GCACGAGCTGAGGAGGAGCAAGAGCACTGTATGTTGGAGCTGCGGAGGAGCT 1133
Db 1133
```

```
Db 721 GCAGCAGCTGGAGGAGCAGAGAAAGCACTGTATGTTTGAAGCTGCGGAGGAGT 780
Qy 1134 GGAGAGTGGCAAGTCTGTGGGAAGTCTTGGCCATCAATGCCAACAACATGTCTCTG 1193
Db 781 GGAGAGTGGCAAGTCTGTGGGAAGTCTTGGCCATCAATGCCAACAACATGTCTCTG 840
Qy 1194 TGCTTGTGCGCGAGCCCCAGGGGCTCTCCCCGGCTGCCCCACTTGGGAGACGGGTCTTC 1253
Db 841 TGCTTGTGCGCGAGCCCCAGGGGCTCTCCCCGGCTGCCCCACTTGGGAGACGGGTCTTC 900
Qy 1254 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 1313
Db 901 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 960
Qy 1314 CACCAACAGCAGGACACAG 1332
Db 961 CACCAACAGCAGGACACAG 979

RESULT 12
US-10-631-958-1
; Sequence 1, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; PRIOR FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 979
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-958-1

Query Match 55.9%; Score 903; DB 19; Length 979;
Best Local Similarity 96.6%; Pred. No. 1.7e-268;
Matches 946; Conservative 0; Mismatches 0; Indels 33; Gaps 1;

Qy 387 ACCAAGCATTTACTGGTATTATCAACCCCGTTTGGAGGAAAGCAAGGCAAGCGGAT 446
Db 1 ACCAAGCATTTACTGGTATTATCAACCCCGTTTGGAGGAAAGCAAGGCAAGCGGAT 60
Qy 447 ATATGAAGAAAAGTGGCACCATTGTTCACTTAGCTTCCATCACCATTGATCATCTG- 505
Db 61 ATATGAAGAAAAGTGGCACCATTGTTCACTTAGCTTCCATCACCATTGATCATCTG 120
Qy 506 -----TTACTGAACATGCTAATCAGGCCAAGGA 533
Db 121 TAACAAATTCATGTTAACTATGTAGAGTAATTACTGAACATGCTAATCAGGCCAAGGA 180
Qy 534 GACTCTGTATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGCGCGGAGATGG 593
Db 181 GACTCTGTATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGCGCGGAGATGG 240
Qy 594 TATGTTGACGAGGTGCTGCGCGGTCTGATTTGGAGGACGAGAGGCGCGGGGTGCA 653
Db 241 TATGTTGACGAGGTGCTGCGCGGTCTGATTTGGAGGACGAGAGGCGCGGGGTGCA 300
Qy 654 CCAGAACCAACCCCGGGCTGTGCTGCCAGTAGCCCTCCGGATTGGAATCATTTCCGCG 713
Db 301 CCAGAACCAACCCCGGGCTGTGCTGCCAGTAGCCCTCCGGATTGGAATCATTTCCGCG 360
Qy 714 AGGGTCAACGGAAGTCTGCTGTGTTACTCCACCGTGGGCAACGACGCGAAGCACTCGGC 773
```

```
Db 361 AGGTCACAGCACTGCGTGTGTACTCCACCGTGGCACCAGCGCAGAAAACCTCGGC 420
Qy 774 GCTGCATATCGTTCTTGGGACATCGCTGGCCATGGATGTCTCCTCAGTCCACCAACAACAG 833
Db 421 GCTGCATATCGTTCTTGGGACATCGCTGGCCATGGATGTCTCCTCAGTCCACCAACAACAG 480
Qy 834 CACACTCTCTCGCTACTCCGCTGCTCCGCTGCTGGGCTACGGCTTCTACGGGACATCATCAA 893
Db 481 CACACTCTCTCGCTACTCCGCTGCTCCGCTGCTGGGCTACGGCTTCTACGGGACATCATCAA 540
Qy 894 GGACAGTGAAGAAAACGGTGGTCTTGGCCAGATACGACTTTTTCAGGTTTAAAGAC 953
Db 541 GGACAGTGAAGAAAACGGTGGTCTTGGCCAGATACGACTTTTTCAGGTTTAAAGAC 600
Qy 954 CTTCTCTCTCCCAACACTGCTATGAAGGACAGTGTCTCTCTCCCTGCACAAACACACGGT 1013
Db 601 CTTCTCTCTCCCAACACTGCTATGAAGGACAGTGTCTCTCTCCCTGCACAAACACACGGT 660
Qy 1014 GGGATCTCCAAGGATAGGAAGCCCTCGCGGGCAGGATGCTTTTTCAGGCAAAAGCAA 1073
Db 661 GGGATCTCCAAGGATAGGAAGCCCTCGCGGGCAGGATGCTTTTTCAGGCAAAAGCAA 720
Qy 1074 GCAGCAGCTGAGGAGGAGCAGAAAGCACTGTATGGTTTGAAGCTGCGGAGGACGT 1133
Db 721 GCAGCAGCTGAGGAGGAGCAGAAAGCACTGTATGGTTTGAAGCTGCGGAGGACGT 780
Qy 1134 GGAGGAGTGGCAAGTCTGCTGCGGAAGTTTCTGGCCATCAATGCCAACAACATGTCTCTG 1193
Db 781 GGAGGAGTGGCAAGTCTGCTGCGGAAGTTTCTGGCCATCAATGCCAACAACATGTCTCTG 840
Qy 1194 TGCTTGTGCGCGGAGCCCCAGGGGCTCTCCCCGGCTGCCACATTTGGGAGACGGGTCTTC 1253
Db 841 TGCTTGTGCGCGGAGCCCCAGGGGCTCTCCCCGGCTGCCACATTTGGGAGACGGGTCTTC 900
Qy 1254 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 1313
Db 901 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTTCTCATCAGGCA 960
Qy 1314 CACCAACAGCAGGACACAG 1332
Db 961 CACCAACAGCAGGACACAG 979

RESULT 13
US-10-635-247
; Sequence 247, Application US/10115635
; Publication No. US20040137434A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: 797CON
; CURRENT APPLICATION NUMBER: US/10/115,635
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 09/714,936
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 362
; SOFTWARE: PE_FL_genes Version 2.0
; SEQ ID NO 247
; LENGTH: 817
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (90)..(662)
US-10-115-635-247

Query Match      35.3%; Score 569; DB 19; Length 817;
Best Local Similarity 100.0%; Pred. No. 4.4e-165;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGGACCGGGGGCGGAGCGCGTGCATCCGTGCTGTGGGTGAAGCAGCAGCGC 60
   |||||
Db 90 ATGGGGGGACCGGGGGCGGAGCGCGTGCATCCGTGCTGTGGGTGAAGCAGCAGCGC 149

QY 61 TCGCCGCTGAGCCTGGAGCCCGCGGGCTCTGCTGCGCTGTGGTGGCGAGCCCGGGGCC 120
   |||||
Db 150 TCGCCGCTGAGCCTGGAGCCCGCGGGCTCTGCTGCGCTGTGGTGGCGAGCCCGGGGCC 209

QY 121 GGAGCGGGCGCCCGGGCGGATCGCTCTGTGCTGTATCTGAGATCATCGCCGTT 180
   |||||
Db 210 GGAGCGGGCGCCCGGGCGGATCGCTCTGTGCTGTATCTGAGATCATCGCCGTT 269

QY 181 GAGGAAAACAGAGCTTCAACGGGAAACATCAAGGCAGTGGAAAATGCAGAAAATGAAAAG 240
   |||||
Db 270 GAGGAAAACAGAGCTTCAACGGGAAACATCAAGGCAGTGGAAAATGCAGAAAATGAAAAG 329

QY 241 CCTTACGCTTTTACAGTTCACTGTGTAAAGAGAGCAGCAGCGCACCGCTGGAAGTGGCG 300
   |||||
Db 330 CCTTACGCTTTTACAGTTCACTGTGTAAAGAGAGCAGCAGCGCACCGCTGGAAGTGGCG 389

QY 301 CAGGTGACTTTCTGTGTCCAGAGAGCAGCTGTCTCACTTGTGCTGCAGACCCCTGGCG 360
   |||||
Db 390 CAGGTGACTTTCTGTGTCCAGAGAGCAGCTGTCTCACTTGTGCTGCAGACCCCTGGCG 449

QY 361 GAGTGTCTGGAGAAGCTGACGTCCAGACCAAGAGCAATTTACTGGTATTTTCAACCCGTT 420
   |||||
Db 450 GAGTGTCTGGAGAAGCTGACGTCCAGACCAAGAGCAATTTACTGGTATTTTCAACCCGTT 509

QY 421 GGAGAAAAGGACAAGCGGAGCGATATATGAAGAAAAGTGGCACCACTGTTCACCTTA 480
   |||||
Db 510 GGAGAAAAGGACAAGCGGAGCGATATATGAAGAAAAGTGGCACCACTGTTCACCTTA 569

QY 481 GCCTCATCACCACCTGACATCATCTGTACTGAACATGCTAATCAGGCCAAGAGACTCTG 540
   |||||
Db 570 GCCTCATCACCACCTGACATCATCTGTACTGAACATGCTAATCAGGCCAAGAGACTCTG 629

QY 541 TATGAGATTAAACATAGACAAATACGACGG 569
   |||||
Db 630 TATGAGATTAAACATAGACAAATACGACGG 658
```

```
RESULT 14
US-09-969-896-8
; Sequence 8, Application US/09969896
; Publication No. US20030125533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-8
```

```
Query Match      29.4%; Score 475; DB 10; Length 550;
Best Local Similarity 100.0%; Pred. No. 4.8e-136;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGGACCGGGGGCGGAGCGCGTGCATCCGTGCTGTGGGTGAAGCAGCAGCGC 60
   |||||
Db 76 ATGGGGGGACCGGGGGCGGAGCGCGTGCATCCGTGCTGTGGGTGAAGCAGCAGCGC 135

QY 61 TGCSCCGTGAACCTGGAGCCCGCGGGCTCTGCTGCGCTGTGGTGGCGAGCCCGGGGCC 120
   |||||
Db 136 TGCSCCGTGAACCTGGAGCCCGCGGGCTCTGCTGCGCTGTGGTGGCGAGCCCGGGGCC 195

QY 121 GGAGCGGGCGCCCGGGCGGATCGCTCTGTGCTGTATCTGAGATCATCGCCGTT 180
   |||||
Db 196 GGAGCGGGCGCCCGGGCGGATCGCTCTGTGCTGTATCTGAGATCATCGCCGTT 255

QY 181 GAGGAAAACAGAGCTTCAACGGGAAACATCAAGGCAGTGGAAAATGCAGAAAATGAAAAG 240
   |||||
Db 256 GAGGAAAACAGAGCTTCAACGGGAAACATCAAGGCAGTGGAAAATGCAGAAAATGAAAAG 315

QY 241 CCTTACGCTTTTACAGTTCACTGTGTAAAGAGAGCAGCAGCGCACCGCTGGAAGTGGCG 300
   |||||
Db 316 CCTTACGCTTTTACAGTTCACTGTGTAAAGAGAGCAGCAGCGCACCGCTGGAAGTGGCG 375

QY 301 CAGGTGACTTTCTGTGTCCAGAGAGCAGCTGTCTCACTTGTGCTGCAGACCCCTGGCG 360
   |||||
Db 376 CAGGTGACTTTCTGTGTCCAGAGAGCAGCTGTCTCACTTGTGCTGCAGACCCCTGGCG 435

QY 361 GAGTGTCTGGAGAAGCTGACGTCCAGACCAAGAGCAATTTACTGGTATTTTCAACCCGTT 420
   |||||
Db 436 GAGTGTCTGGAGAAGCTGACGTCCAGACCAAGAGCAATTTACTGGTATTTTCAACCCGTT 495

QY 421 GGAGAAAAGGACAAGCGGAGCGATATATGAAGAAAAGTGGCACCACTGTTC 475
   |||||
Db 496 GGAGAAAAGGACAAGCGGAGCGATATATGAAGAAAAGTGGCACCACTGTTC 550
```

```
RESULT 15
US-10-631-958-8
; Sequence 8, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-958-8
```

```
Query Match      29.4%; Score 475; DB 19; Length 550;
Best Local Similarity 100.0%; Pred. No. 4.8e-136;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGGACCGGGGGCGGAGCGCGTGCATCCGTGCTGTGGGTGAAGCAGCAGCGC 60
   |||||
Db 76 ATGGGGGGACCGGGGGCGGAGCGCGTGCATCCGTGCTGTGGGTGAAGCAGCAGCGC 135

QY 61 TGCSCCGTGAACCTGGAGCCCGCGGGCTCTGCTGCGCTGTGGTGGCGAGCCCGGGGCC 120
   |||||
```



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 03:56:01 ; Search time 17.8442 Seconds  
(without alignments)  
1363.781 Million cell updates/sec

Title: , US-10-631-958-2  
Perfect score: 1717  
Sequence: 1 PKHLVFINPFGKQKRI.....KCSRNFURFLIRHTNQDQ 326

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A-COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B-COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A-COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B-COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCITUS-COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	613	35.7	687	4	US-09-270-767-45874
2	566	33.0	359	4	US-09-270-767-46720
3	282	16.4	490	4	US-10-053-510-19
4	270.5	15.8	524	4	US-10-053-510-20
5	243	14.2	618	4	US-09-970-516-4
6	243	14.2	618	4	US-09-817-676A-14
7	236	13.7	617	4	US-09-817-676A-12
8	228	13.3	384	4	US-09-949-016-7026
9	228	13.3	384	4	US-09-796-487-3
10	227.5	13.2	384	4	US-09-970-516-2
11	227.5	13.2	384	4	US-09-959-897-2
12	226	13.2	373	4	US-09-796-487-5
13	226	13.2	381	4	US-09-796-487-1
14	226	13.2	381	4	US-09-796-487-4
15	226	13.2	382	4	US-09-970-516-6
16	226	13.2	388	4	US-09-817-676A-15
17	226	13.2	388	4	US-09-796-487-2
18	223.5	13.0	368	4	US-10-053-510-21
19	220.5	12.8	424	4	US-09-796-487-8
20	219.5	12.8	392	4	US-09-796-487-6
21	216	12.6	536	4	US-09-248-796A-15859
22	193	11.2	294	4	US-09-796-487-9
23	181.5	10.6	403	4	US-09-796-487-7
24	172	10.0	299	4	US-09-270-767-61405
25	143	8.3	312	4	US-09-949-016-9811
26	132	7.7	313	4	US-09-107-532A-7154
27	131.5	7.7	372	4	US-09-603-208A-264

28	129.5	7.5	119	4	US-09-205-258-788	Sequence 788, App
29	125	7.3	294	4	US-09-583-110-3740	Sequence 3740, Ap
30	125	7.3	315	4	US-09-107-433-3121	Sequence 3121, Ap
31	118.5	6.9	283	4	US-09-634-238-373	Sequence 373, App
32	117	6.8	160	4	US-09-270-767-62339	Sequence 62339, A
33	117	6.8	304	4	US-09-134-000C-5366	Sequence 5366, Ap
34	110	6.4	303	4	US-09-107-532A-5133	Sequence 5133, Ap
35	109	6.3	325	4	US-09-107-532A-6786	Sequence 6786, Ap
36	108	6.3	293	4	US-09-205-258-328	Sequence 328, App
37	106	6.2	316	4	US-09-710-279-1350	Sequence 1350, Ap
38	106	6.2	324	3	US-09-134-000C-4774	Sequence 4774, Ap
39	105	6.1	345	4	US-09-134-000C-3813	Sequence 3813, Ap
40	104	6.1	63	4	US-09-959-897-45	Sequence 45, Appl
41	102	5.9	64	4	US-09-959-897-37	Sequence 37, Appl
42	92	5.4	65	4	US-09-959-897-53	Sequence 53, Appl
43	90.5	5.3	64	4	US-09-959-897-29	Sequence 29, Appl
44	89.5	5.2	323	4	US-09-107-532A-5598	Sequence 5598, Ap
45	89	5.2	297	4	US-09-134-000C-3616	Sequence 3616, Ap

ALIGNMENTS

RESULT 1

US-09-270-767-45874  
; Sequence 45874, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 45874  
; LENGTH: 687  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-45874

Query Match	35.7%	Score 613;	DB 4;	Length 687;
Best Local Similarity	35.2%	Pred. No. 1.3e-59;		
Matches 134;	Conservative 59;	Mismatches 108;	Indels 80;	Gaps 7;
QY	2	KHLVFINPFGKQKRIYERKVAFLFTLASITTDIIIGNKFYVNYVEVITEHANQAKET	61	
DB	204	RLLVFINPYGGRKAGATYERHVRPIQLAGVATCI-----TTORANQVKDI	252	
QY	62	LYEINIDKYDGVCGDGMFSEVLHGLIGRTQSGAGVDQNHPRAVLPSSLRIGIIPAG	121	
DB	253	LLSHDLGYDVAVCVGGDGTVAEVINGLIFRQMRGLDQRPPIPRP-ALPVGVIIPAG	311	
QY	122	STDCVCYSTGTSDAETSAHLHVGDLSLMDVSSVHHNSTLLRYSVSLGFGYGDIIKD	181	
DB	312	STDTIYSSMHGTADVRTAAIHVILGQHRGLDVCVSNQCSLLRFCAVSLVSLGDLGVAQ	371	
QY	182	SEKKRWLGLARYDFSGKLTFLSHHCYECTVSFL-----PAQHTVGSPRD-----	225	
DB	372	SENYRWMPRIYEGVGVKAFLLNRCYDAELRWLBEPLDILLTTPLEDIFQSPDVSCLGES	431	
QY	226	-RKPCRACFCVC-----ROSKQOLEEOKKALYGLEAAEDVE-----	261	
DB	432	VPSVCYANCORCSFASSIQEQRSLFIQESKEA-----ERNQOVETEDSHLAASEAALLR	487	
QY	262	-----EMQVVCCKFLAINATNMSCACRRSPRGSLSPAHLGD	297	
DB	488	PRPRPGNURLPTGSISSMRNLGNQWKYVRGNFFMICGANTTCACARSPNGISRYSLHGD	547	
QY	298	GSSDLILIRKCSRNFNRLFLI	318	
DB	548	GCLDILLVKKTSLNNVRFLI	568	

```

Query Match      16.4%; Score 282; DB 4; Length 490;
Best Local Similarity 25.0%; Pred.No.1.3e-22;
Matches 104; Conservative 52; Mismatches 134; Indels 126; Gaps 14;

QY    2   KHLVFINPFGGKGOKRIERYKVAFLFTLASITTTDIIGNKFYVNVYEVITEHANOAKE- 60
DB    38   KQLLILLNPKSGSGKRELFOKQVAPLLTEAEVOYDL-----QITTHPOYAKEF 86

QY    61  --TLYEINIDKYDGIVCVGSDGMFSVLHGILIGRTORSAGVDQNHPRAVLVPSSLRIGII 118
DB    87   VTRERDL-LTRYSGIIVVAGDGUFYEYINGLMER-----MDWRRAC--RELPIGLII 134

QY    119  PAGSTDCV-----CVSVTGTSDAETSALHIUVGDSLAMDVSVVH---HNSLTLLRYSVS 168
DB    135  PCDSGNGLAKSAVHAHCNEPYEPKIPILHALUTCWAGKSTPMDVVRVELATRDKHFWMSYFL 194

QY    169  LLGYGFYGDIIKDSEKKRWLGRLARYDFSGKLTLFSLHHCYEGTVSF----- 213
DB    195  SVCGMLIADIDIESERLSIGAQRFTLWAIKRLIGLSYKGRYSYLIGKGKEPPVEAAR 255

QY    214  -LPAQHTVGSVRPRDKPCRCAGCFVCRSKSQOLEEEOKKALYLEAAEDV-----E 261
DB    255  ELPAESTAAGIRSSLPINAGEF----HDLPEEBEGAULDEGFADAI SLDRSVYROHAD 310

QY    262  EQWQVC-----GKFLAINATNNMCACR-----RSPRG----- 288
DB    311  SWSHAMSRRTYAYSLGSPSRNRSMISORIEAANAFAERPVTGTIPLQMPLLSSD 370

QY    289  -----LSPAHLHDGGSDLLILRK-CSRFNFLAPLI 318
DB    371  GWICEDGDFVMVHAAVYTHLSSDVFPAESRLDGLIYLVIIIRRGVSRHQLLNFML 426

RESULT 4
US-10-053-510-20
; Sequence 20, Application US/10053510
; Patent No. 6830881
; GENERAL INFORMATION:
; APPLICANT: Saba, Julie D.
; APPLICANT: Fyrst, Henrik
; TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
; TITLE OF INVENTION: POLYNUCLEOTIDES AND MODULATING AGENTS AND
; METHOD OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 200116.402C2
; CURRENT APPLICATION NUMBER: US/10/053,510
; CURRENT FILING DATE: 2002-01-17
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-053-510-20

Query Match      15.8%; Score 270.5; DB 4; Length 524;
Best Local Similarity 25.2%; Pred.No.2.8e-21;
Matches 104; Conservative 54; Mismatches 124; Indels 131; Gaps 16;

QY    2   KHLVFINPFGGKGOKRIERYKVAFLFTLASITTTDIIGNKFYVNVYEVITEHANOAKE 61
DB    43   RRVLVLLNPKSGSGDAREVENFMHVTPVLNEAVPYDL-----YVTKSNFALEF 91

QY    62   LYEINIDKYDGIVCVGSDGMFSVLHGILIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAG 121
DB    92   LSTRCLDAWCVCVAVGSDGUFHEITVNGLLORQ-----DWAH-----VLPHLALGIIPC 140

QY    122  STDCV-----CVS--TVGTSDAETSALHIVGDSLAMDVSSVHHNSTLLRYSVSLLYGY 173
DB    141  SGNGLAARSIAHCYNKPVIQ-----AALTVISGRSSPM DVTVQLQRSU-YPSFLSIGWG 193

QY    174  FYGDIINKDSEKKRWLGRLARYDFSGKLTLFSLHHCYEGTVSFLPAQHTVGSPRD----- 225
DB    194  LISDVDIESIRIMIGYORTFTVTLBYLVNLRTYNGRISVLLTDHEVSSVSHSATGYAAOR 253

```

```
QY 226 RKPCRACGF-----VCRSQKQLEEE-----QK 248
Db 254 RMQSRSCNTHIDMLNGPAPIYHSSAEYLPQEFADVISLETSINQSRCDSWLGGSR 313
QY 249 KALY-----GLEAAE-----DVEE-----MQVWCG 268
Db 314 RSFYYSISBSIYHSLADESEFAGLAAASLENQQNYGPASELPDLNEPSEDOQWLVSE 373
QY 269 KFLAINA-----TNMSCACRRSPRGLGPAAHLGDSGSDLLIRK-CRSENFRL 317
Db 374 EFWMHAVYQTHGIDCH-----FAPKAQNDGTIYLILIRAGISRPHLLSEL 421

RESULT 5
US-09-970-516-4
; Sequence 4, Application US/09970516
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; FILE OF INVENTION: polynucleotides encoding sphingosine kinases
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-516-4

Query Match 14.2%; Score 243; DB 4; Length 618;
Best Local Similarity 28.9%; Pred. No. 4.6e-18;
Matches 67; Conservative 43; Mismatches 90; Indels 32; Gaps 4;

QY 1 PKHLVFNPFPGKGQKRIYERKVPAPLFTLASITTDIIGNKFYVYVEVITEHANOAKE 60
Db 144 PPRLLLLNVNPFGRGLAWQCKNHVLPMTISEAGLSFNLIQ-----TERQNHARE 192
QY 61 TLYENIDKYDGIYCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPA 120
Db 193 LVQGLSLEWDGIVTVSGDGLLHEVLNGLDR-----PDWEAVKMP-----VGIILPC 240
QY 121 GSTDCVCYS-----TVGTSDAETSALHIWVGDSLAMDVSSVHHNSTLLRYSVSLIG 171
Db 241 GSGNALAGAVNQHGGEFEPALGLDLLNCSLLLCRGGHPDLLSVTLASGSRCSFSLVA 300
QY 172 YGFYGDIIKDEKKRWGLARYDFSLGKTLFSLHHCYEGTVSFLPAQHTVGP 223
Db 301 WGFVSDVDIQSRFRALGSARFTLTGLVLGLATLHTYGRGLSYLPATVPEPASP 352

RESULT 6
US-09-817-676A-14
; Sequence 14, Application US/09817676A
; Patent No. 6800470
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; FILE OF INVENTION: Expression and Methods of Use Thereof
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817,676A
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/194,318
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-817-676A-14

Query Match 13.7%; Score 236; DB 4; Length 617;
Best Local Similarity 29.7%; Pred. No. 2.8e-17;
Matches 66; Conservative 40; Mismatches 84; Indels 32; Gaps 4;

QY 4 LLVFINPFPGKGQKRIYERKVPAPLFTLASITTDIIGNKFYVYVEVITEHANOAKETLY 63
Db 148 LLLVNPFGGRGLAWQRCMDHVPWPMISEAGLSFNLIQ-----TERQNHARELVQ 196
QY 64 EINTDKYDGIYCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPAGST 123
Db 197 GLSLSEWEGIVTVSGDGLLYEVLNGLDR-----PDWEADAVRMP-----IGVLPCGSG 244
QY 124 DVCV-----YSTVGTSDAETSALHIWVGDSLAMDVSSVHHNSTLLRYSVSLIGYGF 174
Db 245 NALAGAVSHHGGFEQVGVVDLLNCSLLLCRGGHPDLLSVTLASGSRCSFSLVANGF 304
QY 175 YGDIKDEKKRWGLARYDFSLGKTLFSLHHCYEGTVSFLPA 216
Db 305 LSDVDIHSERFRALGSARFTLTGAVIGLASLHTYGRGLSYLPA 346

RESULT 8
US-09-949-016-7026
; Sequence 7026, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

```
; ORGANISM: Homo sapiens
US-09-817-676A-14

Query Match 14.2%; Score 243; DB 4; Length 618;
Best Local Similarity 28.9%; Pred. No. 4.6e-18;
Matches 67; Conservative 43; Mismatches 90; Indels 32; Gaps 4;

QY 1 PKHLVFNPFPGKGQKRIYERKVPAPLFTLASITTDIIGNKFYVYVEVITEHANOAKE 60
Db 144 PPRLLLLNVNPFGRGLAWQCKNHVLPMTISEAGLSFNLIQ-----TERQNHARE 192
QY 61 TLYENIDKYDGIYCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPA 120
Db 193 LVQGLSLEWDGIVTVSGDGLLHEVLNGLDR-----PDWEAVKMP-----VGIILPC 240
QY 121 GSTDCVCYS-----TVGTSDAETSALHIWVGDSLAMDVSSVHHNSTLLRYSVSLIG 171
Db 241 GSGNALAGAVNQHGGEFEPALGLDLLNCSLLLCRGGHPDLLSVTLASGSRCSFSLVA 300
QY 172 YGFYGDIIKDEKKRWGLARYDFSLGKTLFSLHHCYEGTVSFLPAQHTVGP 223
Db 301 WGFVSDVDIQSRFRALGSARFTLTGLVLGLATLHTYGRGLSYLPATVPEPASP 352

RESULT 7
US-09-817-676A-12
; Sequence 12, Application US/09817676A
; Patent No. 6800470
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; FILE OF INVENTION: Expression and Methods of Use Thereof
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817,676A
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/194,318
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-817-676A-12

Query Match 13.7%; Score 236; DB 4; Length 617;
Best Local Similarity 29.7%; Pred. No. 2.8e-17;
Matches 66; Conservative 40; Mismatches 84; Indels 32; Gaps 4;

QY 4 LLVFINPFPGKGQKRIYERKVPAPLFTLASITTDIIGNKFYVYVEVITEHANOAKETLY 63
Db 148 LLLVNPFGGRGLAWQRCMDHVPWPMISEAGLSFNLIQ-----TERQNHARELVQ 196
QY 64 EINTDKYDGIYCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPAGST 123
Db 197 GLSLSEWEGIVTVSGDGLLYEVLNGLDR-----PDWEADAVRMP-----IGVLPCGSG 244
QY 124 DVCV-----YSTVGTSDAETSALHIWVGDSLAMDVSSVHHNSTLLRYSVSLIGYGF 174
Db 245 NALAGAVSHHGGFEQVGVVDLLNCSLLLCRGGHPDLLSVTLASGSRCSFSLVANGF 304
QY 175 YGDIKDEKKRWGLARYDFSLGKTLFSLHHCYEGTVSFLPA 216
Db 305 LSDVDIHSERFRALGSARFTLTGAVIGLASLHTYGRGLSYLPA 346

RESULT 8
US-09-949-016-7026
; Sequence 7026, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```



```
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7026
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7026

Query Match      13.3%; Score 228; DB 4; Length 384;
Best Local Similarity 25.9%; Pred. No. 1e-16;
Matches 88; Conservative 53; Mismatches 137; Indels 62; Gaps 13;

Qy 1 PKHLVFINPFGGKGQKRIYERKVAFLFTLASITTDIIGNKFVYVYVEVITEHANOAKE 60
Db 14 PCRVLLNPRGGKGKALQFRSHVQPLLAEEI-----SFTLMLTERRNHARE 62

Qy 61 TLYEINDKYDGVGVGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPA 120
Db 63 LVRSEELGRDALVVMGSDGLMHVNGLMERPDPWETAIQ-----PLCSLIPA 110

Qy 121 GSTDCVC-----YSTVGTSDAETSALHIVVGDLSA-MDVSSVHHNSTLLRYSVSLLG 171
Db 111 GSGNALAASLNHYAGYEQVTNEDLLTCTLLCRLLSPMILLSLHTASGLRFLSVLSLA 170

Qy 172 YGFYGDIIKDEKKRWGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRA 231
Db 171 WGFIADVLESEKRYRLGEMRFTLTGTLRLAALRTYRGLAYLPVGR-VGSKTPASP--- 226

Qy 232 GCFVCRSQKQLEEBEQKALVGLEAAEDV-EWQVVCVK-----FLAINATMNSCACRRSP 286
Db 227 --VVVQOGPVDH-----LVPLE--EPVPSHWTMVPDEDFVLILALHSLGSEMFAAP 276

Qy 287 RGLSPAHLGDSGLDILIRK-CSRFNFRFLI-----RH 320
Db 277 MRCRA-----GVMLFVVRAGVSRAMLLRFLAMEKGRH 311

RESULT 9
US-09-796-487-3
; Sequence 3, Application US/09796487
; Patent No. 6830916
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
; FILE REFERENCE: 07320001aa (2033957-0001)
; CURRENT APPLICATION NUMBER: US/09/796,487
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,532
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: US 09/530,868
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(384)
; OTHER INFORMATION: SEQ ID NO 3 is the peptide sequence of hSPHK1 in Fig. 3, correspond
; OTHER INFORMATION: nding to amino acid residue 1 to 384 of Homo sapiens SPHK-1 of Ge
```

```
; OTHER INFORMATION: nBank sequence Accession Number AAF73423.
; PUBLICATION INFORMATION:
; AUTHORS: Nava et al.
; TITLE: Functional characterization of human sphingosine kinase-1
; JOURNAL: FEBS Lett.
; VOLUME: 473
; ISSUE: 1
; PAGES: 81-84
; DATE: 2000
; DATABASE ACCESSION NUMBER: AAF73423
; DATABASE ENTRY DATE: 2000-06-01
; RELEVANT RESIDUES: (1)..(384)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AAF73423
; DATABASE ENTRY DATE: 2000-06-01
; RELEVANT RESIDUES: (1)..(384)
US-09-796-487-3

Query Match      13.3%; Score 228; DB 4; Length 384;
Best Local Similarity 25.9%; Pred. No. 1e-16;
Matches 88; Conservative 53; Mismatches 137; Indels 62; Gaps 13;

Qy 1 PKHLVFINPFGGKGQKRIYERKVAFLFTLASITTDIIGNKFVYVYVEVITEHANOAKE 60
Db 14 PCRVLLNPRGGKGKALQFRSHVQPLLAEEI-----SFTLMLTERRNHARE 62

Qy 61 TLYEINDKYDGVGVGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPA 120
Db 63 LVRSEELGRDALVVMGSDGLMHVNGLMERPDPWETAIQ-----PLCSLIPA 110

Qy 121 GSTDCVC-----YSTVGTSDAETSALHIVVGDLSA-MDVSSVHHNSTLLRYSVSLLG 171
Db 111 GSGNALAASLNHYAGYEQVTNEDLLTCTLLCRLLSPMILLSLHTASGLRFLSVLSLA 170

Qy 172 YGFYGDIIKDEKKRWGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRA 231
Db 171 WGFIADVLESEKRYRLGEMRFTLTGTLRLAALRTYRGLAYLPVGR-VGSKTPASP--- 226

Qy 232 GCFVCRSQKQLEEBEQKALVGLEAAEDV-EWQVVCVK-----FLAINATMNSCACRRSP 286
Db 227 --VVVQOGPVDH-----LVPLE--EPVPSHWTMVPDEDFVLILALHSLGSEMFAAP 276

Qy 287 RGLSPAHLGDSGLDILIRK-CSRFNFRFLI-----RH 320
Db 277 MRCRA-----GVMLFVVRAGVSRAMLLRFLAMEKGRH 311

RESULT 10
US-09-796-516-2
; Sequence 2, Application US/09970516
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; TITLE OF INVENTION: polynucleotides encoding sphingosine kinases
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-516-2

Query Match      13.2%; Score 227.5; DB 4; Length 384;
Best Local Similarity 27.0%; Pred. No. 1.2e-16;
Matches 64; Conservative 43; Mismatches 97; Indels 33; Gaps 5;

Qy 1 PKHLVFINPFGGKGQKRIYERKVAFLFTLASITTDIIGNKFVYVYVEVITEHANOAKE 60
Db 14 PCRVLLNPRGGKGKALQFRSHVQPLLAEEI-----SFTLMLTERRNHARE 62
```

Qy	61	TLYEINIDKYD	IGV	ICV	GGD	GFSE	VHGL	IGT	QRS	AGV	DQNH	PR	AVL	VP	SSL	RIG	IT	PA	120								
Db	63	LVRSEELGR	WDAL	VMS	GDGL	MHEV	VYV	GM	ER	PK	-----	PL	CS	LP	110												
Qy	121	GSTD	CVC	-----	Y	STV	GT	DA	ET	SA	LH	I	V	GD	SLA	-	MD	VSS	VH	HN	ST	LR	YS	VL	IG	171	
Db	111	GS	G	N	A	A	S	L	N	H	Y	A	G	E	Q	V	T	N	E	D	L	L	T	N	C	T	170
Qy	172	YGF	Y	G	D	I	I	K	O	E	K	R	W	L	G	I	A	R	Y	D	F	S	G	L	T	F	228
Db	171	WG	F	A	D	V	D	L	E	S	K	T	R	L	G	E	M	F	T	L	G	T	F	L	R	A	226

## RESULT 11

```

US-09-953-897-2
; Sequence 2, Application US/09959897
; Patent No. 6730480
; GENERAL INFORMATION:
; APPLICANT: PITSON, Stuart M
; APPLICANT: Brian, WATTENBERG W
; APPLICANT: Fu, Xia
; APPLICANT: Richard, D'ANDREA J
; APPLICANT: Jennifer, BAMBLE R
; APPLICANT: Mathew, VADAS A
; TITLE OF INVENTION: SPHINGOSINE KINASE ENZYME
; FILE REFERENCE: PITSON=1
; CURRENT APPLICATION NUMBER: US/09/959,897
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: PCT/AU00/00457
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: AU PQ 0339
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: AU PQ 1504
; PRIOR FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-959-897-2

```

Query Match	13.2%	Score	227.5	DB	4	Length	384
Best Local Similarity	27.0%	Pred. No.	1.2e-16				
Matches	64	Conservative	43	Mismatches	97	Indels	33
						Gaps	5
Qy	1	PKHLLVFINPPFGGQGGKRIYERKVAFLFTLASITTTDIIGNKFYVYVEVITEHANQAKE	60				
Db	14	PCRVLLINPPFGGKGKALQLFPSHVQPLLAKEI	62				
Qy	61	TLYEINIDKYDGIYVCVGDGMFSEVILHGLIGRTORSAGVDQNHPRAVLPSPSLRIGIIPA	120				
Db	63	LVRSEELGRWDALVVMGSDGLMHVEVNGLMERPPOWETAIQK	110				
Qy	121	GSTDVCV-----YSTVGTSDAETSAHLHIVCGDSL-ADVSSVHHNSTLLRYSVSLIG	171				
Db	111	GSGNALAAASLNHYAGYEQVTTNEDLLTNTLGLCRLLSPMLLLSHTAGSLRFLSVLSLA	170				
Qy	172	YGFYGDIIKQSEKKRWLGLARYDFSGLKTFLSHHCYEGCTVFLPAQHTVGSFDRKP	228				
Db	171	WGFTADVLESEKRYRLGEMFTLGTFLRLAALRYTGRLAYLPVGR-VGSKTPASP	226				

## RESULT 12

```

RES001 12
US-09-796-487-5
; Sequence 5, Application US/09796487
; Patent No. 6830916
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; TITLE OF INVENTION: Spingosin Kinase, Cloning, Expression and Methods of Use
; FILE REFERENCES: 07320001aa (2033957-0001)
; CURRENT APPLICATION NUMBER: US/09796,487

```

```

; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,532
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: US 09/530,868
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(373)
; OTHER INFORMATION: SEQ ID NO 5 is the peptide sequence of Mouse SPK1a in F
; OTHER INFORMATION: responding to amino acid residue 131 to 504 of SPK1a o
; OTHER INFORMATION: sequence Accession Number AAC61697.
; PUBLIC INFORMATION:
; DATABASE ACCESSION NUMBER: AAC61697
; DATABASE ENTRY DATE: 1998-09-26
; RELEVANT RESIDUES: (132)..(504)
; US-09-796-487-5

Query Match 13.2%; Score 226; DB 4; Length 373;
Best Local Similarity 26.6%; Pred. No. 1.7e-16;
Matches 63; Conservative 44; Mismatches 94; Indels 36; Gaps 5;

Qy 1 PKHLLVFINPFGKGQGKRIYERKVAPLETLASITTTDIIGNKPYVNYVEVITEHANOAKE 60
Db 5 PCRVLLINPQGGKALQLFQSRVQPFLEAEITPKLI-----LTERKNHARE 53
Qy 61 TLYEINIDKYDGVICVCGGDMGFSEVLHGLIGRTQRSAGVDQNHPRAVLVFPSSLRIGIIPA 120
Db 54 LVCAEELGHWDALAVNSGDLGMEVNVGLMERPDWETAIQK-----PLCSLPG 101
Qy 121 GSTDCVC-----YSTVGTSDAETSALHIVGDSLIA-MDYSSVHNSTLLRYSVSLIG 171
Db 102 GSGNALAASVNHVAGYEQVNTNEDLLNCTLLCRRLSPMNLISLTASGLRLYSLVS 161
Qy 172 YGFYGDIIIDSEKKRWGLGARYDFSGLKTFLSHHCYEGVTSFLPAQHTVGSPRDRKP 228
Db 162 WGFVADVLESEKRYRIGRTVGTGTFPFIASLRIVYOGIAVLP-----VGTWASKRP 214

```

RESIT.T 13

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```
; TITLE: Molecular cloning and functional characterization of murine sphingosin
; TITLE: kinase
; JOURNAL: Journal of Biological Chemistry
; VOLUME: 237
; ISSUE: 37
; PAGES: 23722-23728
; DATE: 1998
; DATABASE ACCESSION NUMBER: AAC61697
; DATABASE ENTRY DATE: 1998-09-26
; RELEVANT RESIDUES: (124)..(504)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AAC61697
; DATABASE ENTRY DATE: 1998-09-26
; RELEVANT RESIDUES: (124)..(504)
US-09-796-487-1

Query Match      13.2%; Score 226; DB 4; Length 381;
Best Local Similarity 26.6%; Pred. No. 1.7e-16;
Matches 63; Conservative 44; Mismatches 94; Indels 36; Gaps 5;

Qy 1 PKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIGNKFYVNVVEVITEHANOAKE 60
Db 13 PCRVLLVLPNPGGKGKALQLQFSRVQPFLEAEITFKLI-----LTERKNHARE 61

Qy 61 TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPA 120
Db 62 LVCAEELGHWDALAVMSGDGLMHEVNVGLMERPDWETAIOK-----PLCSLPG 109

Qy 121 GSTDCVC-----YSTVGTSDAETSALHIVVGDLSA-MDVSSVHHNSTLLRYSVSLIG 171
Db 110 GSGNALAASVNHVAGYEQVTNEDLLINCTLLCRRRLSPMNLLSLHTASGLRLYSVLSLS 169

Qy 172 YGFYGDIIKDKSEKKRWLGARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPPDRKP 228
Db 170 WGFVADVLESEKRYRLGEIRFTVGTFFRLASLRIOGQLAYLP-----VGTVASKRP 222

RESULT 14
US-09-796-487-4
; Sequence 4, Application US/09796487
; Patent No. 6830916
; GENERAL INFORMATION:
; APPLICANT: Spiegell, Sarah
; TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
; FILE REFERENCE: 07320001aa (2033957-0001)
; CURRENT APPLICATION NUMBER: US/09/796,487
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,532
; PRIOR FILING DATE: 2000-03-03
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(381)
; OTHER INFORMATION: SEQ ID NO 4 is the peptide sequence of mSPHK1a in Figure 3, corre
; OTHER INFORMATION: sponding to amino acid residue 124 to 504 of SPHK1a of GenBank se
; OTHER INFORMATION: quence Accession Number AF068748. SEQ ID NO 4 is equivalent to SE
; OTHER INFORMATION: Q ID NO 1 that is the amino acid sequence of SPHK1a in Figure 1.
; PUBLICATION INFORMATION:
; AUTHORS: Kohama et al.,
; TITLE: Molecular cloning and functional characterization of murine sphingosine
; JOURNAL: Journal of Biological Chemistry
; VOLUME: 273
; ISSUE: 37
; PAGES: 23722-23728
; DATE: 1998
```

```
; DATABASE ACCESSION NUMBER: AAC61697
; DATABASE ENTRY DATE: 1998-09-26
; RELEVANT RESIDUES: (124)..(504)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AAC61697
; DATABASE ENTRY DATE: 1998-09-26
; RELEVANT RESIDUES: (124)..(504)
US-09-796-487-4

Query Match      13.2%; Score 226; DB 4; Length 381;
Best Local Similarity 26.6%; Pred. No. 1.7e-16;
Matches 63; Conservative 44; Mismatches 94; Indels 36; Gaps 5;

Qy 1 PKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIGNKFYVNVVEVITEHANOAKE 60
Db 13 PCRVLLVLPNPGGKGKALQLQFSRVQPFLEAEITFKLI-----LTERKNHARE 61

Qy 61 TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPA 120
Db 62 LVCAEELGHWDALAVMSGDGLMHEVNVGLMERPDWETAIOK-----PLCSLPG 109

Qy 121 GSTDCVC-----YSTVGTSDAETSALHIVVGDLSA-MDVSSVHHNSTLLRYSVSLIG 171
Db 110 GSGNALAASVNHVAGYEQVTNEDLLINCTLLCRRRLSPMNLLSLHTASGLRLYSVLSLS 169

Qy 172 YGFYGDIIKDKSEKKRWLGARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPPDRKP 228
Db 170 WGFVADVLESEKRYRLGEIRFTVGTFFRLASLRIOGQLAYLP-----VGTVASKRP 222

RESULT 15
US-09-970-516-6
; Sequence 6, Application US/09970516
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; FILE REFERENCE: polynucleotides encoding sphingosine kinases
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 6
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-970-516-6

Query Match      13.2%; Score 226; DB 4; Length 382;
Best Local Similarity 26.6%; Pred. No. 1.7e-16;
Matches 63; Conservative 44; Mismatches 94; Indels 36; Gaps 5;

Qy 1 PKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIGNKFYVNVVEVITEHANOAKE 60
Db 14 PCRVLLVLPNPGGKGKALQLQFSRVQPFLEAEITFKLI-----LTERKNHARE 62

Qy 61 TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPA 120
Db 63 LVCAEELGHWDALAVMSGDGLMHEVNVGLMERPDWETAIOK-----PLCSLPG 110

Qy 121 GSTDCVC-----YSTVGTSDAETSALHIVVGDLSA-MDVSSVHHNSTLLRYSVSLIG 171
Db 111 GSGNALAASVNHVAGYEQVTNEDLLINCTLLCRRRLSPMNLLSLHTASGLRLYSVLSLS 170

Qy 172 YGFYGDIIKDKSEKKRWLGARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPPDRKP 228
Db 171 WGFVADVLESEKRYRLGEIRFTVGTFFRLASLRIOGQLAYLP-----VGTVASKRP 223

Search completed: September 3, 2005, 04:10:35
Job time : 19.8442 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 04:09:18 ; Search time 66.5726 Seconds  
(without alignments)  
1893.930 Million cell updates/sec

Title: US-10-631-958-2  
Perfect score: 1717  
Sequence: 1 PKHLVFNPGGKCGKRI.....KCSRNFRLRHTNQDQ 326

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A: Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1717	100.0	326	5	ABb07854 Human sph
2	1717	100.0	471	4	Aae07884 Human sph
3	1640.5	95.5	416	6	ABr56301 Human sph
4	1640.5	95.5	460	3	Aay96059 Human sph
5	1640.5	95.5	481	6	ABr56302 Human sph
6	1640.5	95.5	537	5	ABb07856 Human sph
7	1640.5	95.5	537	6	Ada05680 Human NOV
8	1640.5	95.5	537	8	Adj96664 Human lip
9	1640.5	95.5	537	8	Adm62845 Human NOV
10	1640.5	95.5	562	5	ABb07857 Human sph
11	1632.5	95.1	537	5	Aam49115 Human cer
12	1503.5	87.6	531	8	Adp55248 Human PRO
13	1414.5	82.4	536	8	Adsl1054 Human the
14	1400.5	81.6	746	4	ABg13541 Novel hum
15	1323.5	77.1	727	4	ABg13543 Novel hum
16	1055	61.4	228	3	Aab41822 Human ORF
17	966	56.3	454	4	ABg13544 Novel hum
18	664	38.7	136	4	ABg13540 Novel hum
19	628	36.6	136	4	ABg13542 Novel hum
20	613	35.7	596	4	ABb69669 Drosophil
21	476.5	27.8	182	3	Aab42383 Human ORF
22	442	25.7	532	8	Adq88891 Novel hum
23	303.5	17.7	255	8	Adsl1267 Human the
24	300.5	17.5	190	5	ABp64913 Human pro
25	282	16.4	490	6	ABr82389 D. melano

26	282	16.4	641	4	ABb58465	Drosophil
27	282	16.4	641	6	ABr82392	D. melano
28	273	15.9	144	4	AAe07885	Partial r
29	270.5	15.8	524	6	ABr82390	D. melano
30	270.5	15.8	907	4	ABb57980	Drosophil
31	270.5	15.8	907	6	ABr82393	D. melano
32	249	14.5	618	4	Aau09075	Human sph
33	243	14.2	618	4	AAU09074	Human sph
34	243	14.2	618	5	ABg31587	Human sph
35	243	14.2	618	5	ABb07855	Protein s
36	243	14.2	618	7	ADe38417	Human pro
37	243	14.2	618	8	ADQ15182	Human can
38	243	14.2	654	4	ABU52806	Human sig
39	243	14.2	654	8	ADJ66569	Sphingosi
40	243	14.2	806	4	ABG21144	Novel hum
41	236	13.7	617	4	AAU09073	Human sph
42	233.5	13.6	384	3	AAy96057	Human sph
43	228	13.3	384	5	ABb08089	Human sph
44	228	13.3	384	8	ADQ17905	Human sof
45	227.5	13.2	384	3	AAb18659	A human r

ALIGNMENTS

RESULT 1

ABb07854  
ID ABb07854 standard; protein; 326 AA.

XX AC ABb07854;

XX AC ABb07854;

XX 03-JUL-2002 (first entry)

XX Human sphingosine kinase-like protein.

XX Human sphingosine kinase-like protein.

KW Human sphingosine kinase-like protein; intracellular signalling;

KW cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;

KW autoimmune disease; rheumatoid arthritis; Parkinson's disease.

XX Homo sapiens.

XX WO200228906-A2.

XX PD 11-APR-2002.

XX PF 05-OCT-2001; 2001WO-EP011516.

XX PR 06-OCT-2000; 2000US-0238005P.

XX PR 23-AUG-2001; 2001US-0314113P.

XX (FARB ) BAYER AG.

XX Kossida S, Encinas J;

XX WPI; 2002-340094/37.

XX DR N-PSDB; ABL40822.

XX New reagent for modulating the activity of sphingosine kinase-like protein polypeptide or polynucleotide and treating cancer, asthma, allergy, an autoimmune disease, or a central or peripheral nervous system disorder.

XX Claim 25; Fig 2; 120pp; English.

XX The invention relates to a human sphingosine kinase-like protein. The polypeptide can be expressed by standard recombinant methodology. The sphingosine kinase-like protein and gene can be used to regulate intracellular signalling and consequently cell proliferation and apoptosis. Such regulation is useful for treating cancer, allergies (e.g. asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and peripheral nervous system disorders (e.g. Parkinson's disease). The present sequence represents the human sphingosine kinase-like protein



useful in e.g. chromosome and gene mapping, and detecting inflammation or disease associated with abnormal levels of sphingosine kinase expression.

Disclosure; Fig 9; 81pp; English.

The present sequence is that of human sphingosine kinase C (SKC), an enzyme that phosphorylates sphingosine to form sphingosine 1-phosphate. The sequence was deduced from that of a polynucleotide (see AAA50510) isolated from an HeLa cDNA library. The invention provides polynucleotides (see AAA50508-10) and polypeptides (see AAY96057-59) for the human sphingosine kinase (SK) homologues SKA, SKB and SKC. The polypeptides can be obtained using recombinant DNA methods, and host cells containing expression vectors including SK polynucleotides are used in a claimed method of screening for compounds that inhibit or activate human SK activity. Human SK specific antibodies, inhibitors, ligands or their analogues can be used as bioactive agents to treat inflammation or disease including viral, bacterial or fungal infections, allergic responses, mechanical injury associated with trauma, hereditary diseases, lymphoma or carcinoma, and other conditions with activate the genes of kidney, lung, heart, lymphoid or tissues of the nervous system

Sequence 460 AA;

Query Match 95.5%; Score 1640.5; DB 3; Length 460;  
Best Local Similarity 96.3%; Pred. No. 2.8e-164;  
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1

QY 1 PKHLLVFIPFGGKGOGKRIYERKVAFLFTLASITTDIIIGNKFYVNVVEITEHANQAKE 60  
DB 53 PKHLLVFIPFGGKGOGKRIYERKVAFLFTLASITTDII-----VTEHANQAKE 101

QY 61 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPA 120  
DB 102 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPA 161

QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGFYGDIK 180  
DB 162 GSTDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGFYGDIK 221

QY 181 DSEKKRWGLIARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDKPCRCAGCFVCROSK 240  
DB 222 DSEKKRWGLIARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDKPCRCAGCFVCROSK 281

QY 241 QLEEEQKKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300  
DB 282 QLEEEQKKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 341

QY 301 DLILIRKCSRNFNPLRLIRHTNQDDQ 326  
DB 342 DLILIRKCSRNFNPLRLIRHTNQDDQ 367

RESULT 5  
ABR56302  
ID ABR56302 standard; protein; 481 AA.  
AC ABR56302;  
XX  
XX  
XX  
XX 20-NOV-2003 (first entry)  
DT  
DE Human Sphingosine kinase 4-related protein.  
XX  
XX Human; enzyme; haemostatic; sphingosine kinase 4; SPHK4;  
KW platelet transfusion; platelet stabiliser.  
XX  
XX Homo sapiens.  
XX  
XX WO2003031627-A1.  
PN  
XX  
XX 17-APR-2003.  
PD  
XX  
XX 28-SEP-2001; 2001WO-JP008537.  
XX

PT treatment of sphingosine related disorders.  
XX  
XX Claim 2; Page 29-30; 39pp; Japanese.  
XX  
XX  
XX The present sequence is the protein sequence for human sphingosine kinase 4 (SPHK4). The kinase can be used for the diagnosis and treatment of sphingosine related disorders. The kinase can also be potentially used for controlling toxicity of platelet transfusion and as a platelet stabiliser  
XX  
XX  
XX Sequence 416 AA;  
SQ

Query Match 95.5%; Score 1640.5; DB 6; Length 416;  
Best Local Similarity 96.3%; Pred. No. 2.4e-164;  
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

QY 1 PKHLLVFIPFGGKGOGKRIYERKVAFLFTLASITTDIIIGNKFYVNVVEITEHANQAKE 60  
DB 9 PKHLLVFIPFGGKGOGKRIYERKVAFLFTLASITTDII-----VTEHANQAKE 57

QY 61 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPA 120  
DB 58 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPA 117

QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGFYGDIK 180  
DB 118 GSTDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGFYGDIK 177

QY 181 DSEKKRWGLIARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDKPCRCAGCFVCROSK 240  
DB 178 DSEKKRWGLIARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDKPCRCAGCFVCROSK 237

QY 241 QLEEEQKKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300  
DB 238 QLEEEQKKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 297

QY 301 DLILIRKCSRNFNPLRLIRHTNQDDQ 326  
DB 298 DLILIRKCSRNFNPLRLIRHTNQDDQ 323

RESULT 4  
AAY96059  
ID AAY96059 standard; protein; 460 AA.  
XX  
XX  
XX  
XX AAY96059;  
DT  
DT 05-DEC-2000 (first entry)  
XX  
XX Human sphingosine kinase C.  
DE  
DE  
KW Sphingosine kinase C; SKC; human; drug screening; infection;  
KW antiinflammatory; antiallergic; anticancer; inflammation; allergy;  
KW cancer; therapy; diagnosis.  
XX  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200052173-A2.  
PN  
XX  
XX 08-SEP-2000.  
PD  
XX  
XX 02-MAR-2000; 2000WO-CA000223.  
PF  
PF  
PR 02-MAR-1999; 99US-0122516P.  
XX  
XX (ALIX ) NPS ALLELIX CORP.  
PA  
XX  
XX Munroe D, Gupta A, Falzone GR;  
PI  
PI WPI; 2000-572185/53.  
DR  
DR N-PSDB; AAA50510.  
XX  
XX New human sphingosine kinase A, B and C polynucleotides and polypeptides

```

PR 28-SEP-2001; 2001WO-JP008537.
XX (HOKK-) HOKKAIDO TECHNOLOGY LICENSING OFFICE CO.
PA (CHBI-) CHEM BIOLOGY INST.
XX
XX Igarashi Y, Kihara A;
XX WPI; 2003-354917/33.
XX
XX Platelet derived polypeptides with sphingosine kinase activity for
XX treatment of sphingosine related disorders.
XX
XX Example 4; Fig 4; 39pp; Japanese.
XX
XX The present invention relates to human sphingosine kinase 4 (SPHK4;
XX AB956301). The kinase can be used for the diagnosis and treatment of
XX sphingosine related disorders. The kinase can also be potentially used
XX for controlling toxicity of platelet transusion and as a platelet
XX stabiliser. The present sequence was used to illustrate the invention
XX
XX Sequence 481 AA;
XX
XX Query Match 95.5%; Score 1640.5; DB 6; Length 481;
XX Best Local Similarity 96.3%; Pred. No. 3e-164;
XX Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;
XX
Qy 1 PKHLLVFINPFGGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVNVVEVITEHANOAKE 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 74 PKHLLVFINPFGGKGQKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 122
Qy 61 TLYEINIDKYDGIIVCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPA 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123 TLYEINIDKYDGIIVCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPA 182
Qy 121 GSTDCVCYSTVGTSDAETSALHIVVGDSDLAMDVSVVHNSLTLRYSVSLGFGYGDIIK 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 183 GSTDCVCYSTVGTSDAETSALHIVVGDSDLAMDVSVVHNSLTLRYSVSLGFGYGDIIK 242
Qy 181 DSEKRWGLGLARYDFSGLKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRCAGCFVCQSK 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 243 DSEKRWGLGLARYDFSGLKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRCAGCFVCQSK 302
Qy 241 QOLEEEQKALYGLEAAEDVEEQVCGKFLAINATNMNSCARRSPRGLSPAHLGDGSS 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 303 QOLEEEQKALYGLEAAEDVEEQVCGKFLAINATNMNSCARRSPRGLSPAHLGDGSS 362
Qy 301 DLILIRKCSRFNFLRLIRHTNQDQ 326
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 363 DLILIRKCSRFNFLRLIRHTNQDQ 388
XX
RESULT 6
AB07856
ID AB07856 standard; protein; 537 AA.
XX
XX AC AB07856;
XX
XX
XX 03-JUL-2002 (first entry)
XX
XX Human sphingosine kinase-like protein.
XX
XX Human sphingosine kinase-like protein; intracellular signalling;
XX cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;
XX autoimmune disease; rheumatoid arthritis; Parkinson's disease.
XX
XX Homo sapiens.
XX
XX WO200228906-A2.
XX
XX 11-APR-2002.
XX
XX 05-OCT-2001; 2001WO-EP011516.
XX
XX

```

---

```

PR 06-OCT-2000; 2000US-023800SP.
XX 23-AUG-2001; 2001US-0314113P.
XX (FARB ) BAYER AG.
XX
XX Kossida S, Encinas J;
XX WPI; 2002-340094/37.
XX N-PSDB; ABL40828.
XX
XX New reagent for modulating the activity of sphingosine kinase-like
XX protein polypeptide or polynucleotide and treating cancer, asthma,
XX allergy, an autoimmune disease, or a central or peripheral nervous system
XX disorder.
XX
XX Claim 25; Fig 10; 120pp; English.
XX
XX The invention relates to a human sphingosine kinase-like protein. The
XX polypeptide can be expressed by standard recombinant methodology. The
XX sphingosine kinase-like protein and gene can be used to regulate
XX intracellular signalling and consequently cell proliferation and
XX apoptosis. Such regulation is useful for treating cancer, allergies (e.g.
XX asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and
XX peripheral nervous system disorders (e.g. Parkinson's disease). The
XX present sequence represents the human sphingosine kinase-like protein
XX
XX Sequence 537 AA;
XX
XX Query Match 95.5%; Score 1640.5; DB 5; Length 537;
XX Best Local Similarity 96.3%; Pred. No. 3.5e-164;
XX Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;
XX
Qy 1 PKHLLVFINPFGGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVNVVEVITEHANOAKE 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 130 PKHLLVFINPFGGKGQKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 178
Qy 61 TLYEINIDKYDGIIVCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPA 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 179 TLYEINIDKYDGIIVCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPA 238
Qy 121 GSTDCVCYSTVGTSDAETSALHIVVGDSDLAMDVSVVHNSLTLRYSVSLGFGYGDIIK 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 239 GSTDCVCYSTVGTSDAETSALHIVVGDSDLAMDVSVVHNSLTLRYSVSLGFGYGDIIK 298
Qy 181 DSEKRWGLGLARYDFSGLKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRCAGCFVCQSK 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 299 DSEKRWGLGLARYDFSGLKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRCAGCFVCQSK 358
Qy 241 QOLEEEQKALYGLEAAEDVEEQVCGKFLAINATNMNSCARRSPRGLSPAHLGDGSS 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 359 QOLEEEQKALYGLEAAEDVEEQVCGKFLAINATNMNSCARRSPRGLSPAHLGDGSS 418
Qy 301 DLILIRKCSRFNFLRLIRHTNQDQ 326
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 419 DLILIRKCSRFNFLRLIRHTNQDQ 444
XX
RESULT 7
ADA05680
ID ADA05680 standard; protein; 537 AA.
XX
XX AC ADA05680;
XX
XX 06-NOV-2003 (first entry)
XX
XX Human NOV9a protein SEQ ID NO:40.
XX
XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
XX immunomodulator; cytostatic; nootropic; neuroprotective;
XX antiparkinsonian; antilipemic; gene therapy; human disease;
XX metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX immune disorder; haematopoietic disorder; dyslipidaemia.
XX

```



XX Homo sapiens.  
XX WO2003029424-A2.  
XX 10-APR-2003.  
XX 02-OCT-2002; 2002WO-US031373.  
XX 02-OCT-2001; 2001US-0326483P.  
XX 05-OCT-2001; 2001US-0327435P.  
XX 05-OCT-2001; 2001US-0327435P.  
XX 09-OCT-2001; 2001US-0327917P.  
XX 09-OCT-2001; 2001US-0328029P.  
XX 09-OCT-2001; 2001US-0328044P.  
XX 09-OCT-2001; 2001US-0328056P.  
XX 12-OCT-2001; 2001US-0328056P.  
XX 15-OCT-2001; 2001US-0329414P.  
XX 17-OCT-2001; 2001US-0330142P.  
XX 18-OCT-2001; 2001US-0330309P.  
XX 22-OCT-2001; 2001US-0341058P.  
XX 24-OCT-2001; 2001US-0339266P.  
XX 24-OCT-2001; 2001US-0339266P.  
XX 29-OCT-2001; 2001US-0343629P.  
XX 29-OCT-2001; 2001US-0346357P.  
XX 01-NOV-2001; 2001US-0346357P.  
XX 17-APR-2002; 2002US-0373260P.  
XX 19-APR-2002; 2002US-0373815P.  
XX 19-APR-2002; 2002US-0373817P.  
XX 19-APR-2002; 2002US-0373826P.  
XX 19-APR-2002; 2002US-0373884P.  
XX 22-APR-2002; 2002US-0374977P.  
XX 16-MAY-2002; 2002US-0381037P.  
XX 16-MAY-2002; 2002US-0381038P.  
XX 16-MAY-2002; 2002US-0381042P.  
XX 17-MAY-2002; 2002US-0381642P.  
XX 28-MAY-2002; 2002US-0383656P.  
XX 29-MAY-2002; 2002US-0383831P.  
XX 25-JUN-2002; 2002US-0391355P.  
XX 01-OCT-2002; 2002US-00262511.  
XX (CURA-) CURAGEN CORP.  
XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
PI Patcurajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
PI Ort T, Gorman L, Zehrhusen BD, Anderson DW, Zhong M, Catterton E;  
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
XX WPI; 2003-381626/36.  
XX N-PSDB; ADA05679.  
XX New NOVX polypeptides and nucleic acids, useful for diagnosing,  
PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.  
XX Claim 1; Page 135; 586pp; English.  
XX The present invention describes NOVX proteins, where X can be 1 to 55  
CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
CC described above and a carrier; (2) a kit comprising, in one or more  
CC containers, the composition described above; (3) an isolated nucleic acid  
CC molecule which encodes a NOVX protein of the invention; (4) a vector  
CC comprising the nucleic acid molecule described above; (5) a cell  
CC comprising the above vector; (6) an antibody that immunospecifically  
CC binds to the polypeptide described above; (7) methods for determining the  
CC presence or amount of the above polypeptide or nucleic acid molecule in a  
CC sample; (8) methods for determining the presence of or predisposition to  
CC a disease associated with altered levels of expression of the above  
CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
CC method of identifying an agent that binds to the polypeptide described  
CC above; (10) a method for identifying a potential therapeutic agent for

CC use in treating a pathology that is related to an aberrant expression or  
CC aberrant physiological interactions of the polypeptide; (11) a method of  
CC screening for a modulator of activity or of latency or predisposition to  
CC a pathology associated with the polypeptide; (12) a method for modulating  
CC the activity of the polypeptide described above; (13) methods of treating  
CC or preventing a pathology associated with the above polypeptide in a  
CC mammal; and (14) a method for producing the above polypeptide. NOVX  
CC sequences have antidiabetic, anorectic, antibacterial, virucide,  
CC immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian  
CC and antilipemic activities, and can be used in gene therapy. The  
CC polypeptide is useful in manufacturing a medicament for treating a  
CC syndrome associated with a human disease. The polypeptide or the nucleic  
CC acid molecule may be used to diagnose, treat or prevent metabolic  
CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
CC disease, immune disorders, haematopoietic disorders and various  
CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
CC probes, in chromosome mapping, tissue typing, preventive medicine and  
CC pharmacogenomics. The present sequence represents a human NOVX from the  
CC present invention.  
XX Sequence 537 AA;  
SQ Query Match 95.5%; Score 1640.5; DB 6; Length 537;  
Best Local Similarity 96.3%; Pred No. 3.5e-164;  
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;  
QY 1 PKHLVFINPFGGKQGRKRIYERKVPFLFTLASITTDIIGNKFYVYVEVITEHANQAKE 60  
DB 130 PKHLVFINPFGGKQGRKRIYERKVPFLFTLASITTDI-----VTEHANQAKE 178  
QY 61 TLYEINIDKYGIVCVGGDGMFSEVLHGLIGRTQTSAGVDQNHPRAVLVPSSLRIGIIPA 120  
DB 179 TLYEINIDKYGIVCVGGDGMFSEVLHGLIGRTQTSAGVDQNHPRAVLVPSSLRIGIIPA 238  
QY 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHNSLTLLRYSVSLGFGFYGDITK 180  
DB 239 GSTDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHNSLTLLRYSVSLGFGFYGDITK 298  
QY 181 DSEKKRWLGLARYDFSGLKTFLSHHCYEGTGVSFPAQHTVGSPPDRKPCACFCVCRQSK 240  
DB 299 DSEKKRWLGLARYDFSGLKTFLSHHCYEGTGVSFPAQHTVGSPPDRKPCACFCVCRQSK 358  
QY 241 QOLEBEQKALYGLEAAEDVEEWQVCGKFLAINATNNSCACRRSPRGLSPAHLGDGSS 300  
DB 359 QOLEBEQKALYGLEAAEDVEEWQVCGKFLAINATNNSCACRRSPRGLSPAHLGDGSS 418  
QY 301 DLILIRKCSRFNLFRLIRHTNQDDQ 326  
DB 419 DLILIRKCSRFNLFRLIRHTNQDDQ 444  
RESULT 8  
ADJ96664  
ID ADJ96664 standard; protein; 537 AA.  
XX AC ADJ96664;  
XX DT 06-MAY-2004 (first entry)  
XX DE Human lipid kinase KIAA1646 protein SeqID 121.  
XX KW kinase; human; tyrosine protein kinase; serine/threonine protein kinase;  
XX KW PTK; STK; gene therapy; cancer; immune-related disease;  
XX KW cardiovascular disease; brain; neuronal associated disease; metabolic;  
XX KW inflammatory disorder; cytostatic; neuroprotective; immunomodulator;  
XX KW antinflammatory; enzyme; lipid kinase; KIAA1646.  
XX OS Homo sapiens.  
XX OS 44.  
XX WO2004006838-A2.  
XX

PD	22-JAN-2004.	
XX		XX
PF	15-JUL-2003; 2003WO-US021730.	KW
XX		KW
PR	15-JUL-2002; 2002US-0395632P.	KW
XX		KW
PA	(SUGE-) SUGEN INC.	XX
XX		OS
XX	Whyte D, Manning G, Caenepeel S;	XX
PI		XX
XX		PN
DR	WPI, 2004-122753/12.	XX
DR	N-FSDB; ADJ96598.	XX
XX		PD
PT	New nucleic acid molecule encoding a kinase polypeptide, useful for	XX
PT	preparing a composition for treating diseases or disorders, e.g., cancer,	PF
PT	or neurological, immunological or inflammatory disorders.	XX
PS	Claim 1; SEQ ID NO 121; 366pp; English.	XX
XX		01-OCT-2002; 2002US-00262511.
CC	This invention relates to a novel isolated, enriched or purified nucleic	02-OCT-2001; 2001US-0326483P.
CC	acid molecule that encodes a kinase polypeptide. Specifically, it relates	05-OCT-2001; 2001US-0327435P.
CC	to human tyrosine and serine/threonine protein kinases (PTK's and STK's),	05-OCT-2001; 2001US-0327449P.
CC	as well as protein kinase-like enzymes. The present invention describes	09-OCT-2001; 2001US-0327917P.
CC	screening methods to identify agonists, antagonists and antibodies that	09-OCT-2001; 2001US-0328029P.
CC	can be used to modulate the activity or function of the mammalian kinase	09-OCT-2001; 2001US-0328044P.
CC	enzymes. As such, these compositions can be used for gene therapy	09-OCT-2001; 2001US-0328056P.
CC	purposes to treat diseases or disorders including cancer, immune-related	12-OCT-2001; 2001US-0328849P.
CC	diseases, cardiovascular disease, brain or neuronal associated disease,	15-OCT-2001; 2001US-0329414P.
CC	metabolic and inflammatory disorders. Accordingly, they exhibit	17-OCT-2001; 2001US-0330142P.
CC	cytostatic, neuroprotective, immunomodulator and antiinflammatory	18-OCT-2001; 2001US-0330309P.
CC	activities. This polypeptide sequence is a human kinase protein sequence	22-OCT-2001; 2001US-0341058P.
CC	of the invention.	24-OCT-2001; 2001US-0339266P.
XX		24-OCT-2001; 2001US-0343629P.
SQ	Sequence 537 AA;	29-OCT-2001; 2001US-0349575P.
		01-NOV-2001; 2001US-0346357P.
		17-APR-2002; 2002US-0373260P.
		19-APR-2002; 2002US-0373815P.
		19-APR-2002; 2002US-0373817P.
		19-APR-2002; 2002US-0373826P.
		19-APR-2002; 2002US-0373884P.
		22-APR-2002; 2002US-0374977P.
		16-MAY-2002; 2002US-0381037P.
		16-MAY-2002; 2002US-0381038P.
		16-MAY-2002; 2002US-0381042P.
		17-MAY-2002; 2002US-0381642P.
		28-MAY-2002; 2002US-0383656P.
		29-JUN-2002; 2002US-0383831P.
		25-JUN-2002; 2002US-0391335P.
QY	1 PKHLVFNPPGGKGGRRIYERKVPAPLFTLASITTDIIGNKFYVNVVEVTEHANQAKE 60	(SMIT/) SMITHSON G.
Db	130 PKHLVFNPPGGKGGRRIYERKVPAPLFTLASITTDII-----VTEHANQAKE 178	(MILL/) MILLET I.
QY	61 TLYEINIDYDGI VCVGSDGMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSPSLRIGIIPA 120	FA (PEYM/) PEYMAN J A.
Db	179 TLYEINIDYDGI VCVGSDGMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSPSLRIGIIPA 238	PA (KEKU/) KEKUDA R.
QY	121 GSTDCVCYSTVGTSDAETSALHIVVGDSLAMDVSVHNSHTLLRYSVSLLGYGFGYDIIK 180	PA (JUJJ/) JU J.
Db	239 GSTDCVCYSTVGTSDAETSALHIVVGDSLAMDVSVHNSHTLLRYSVSLLGYGFGYDIIK 298	PA (LILL/) LI L.
QY	181 DSEKKRWGLARYDFSGIKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRAGCFVCRQSK 240	PA (GUOX/) GUO X.
Db	299 DSEKKRWGLARYDFSGIKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRAGCFVCRQSK 358	FA (PATT/) PATTURAJAN M.
QY	241 QOLBEOKKALYGLEAEDVVEQVVCVKFLAINATNNSCACRRSPRGLSPAHLGDGSS 300	FA (SPYT/) SPYTEK K A.
Db	359 QOLBEOKKALYGLEAEDVVEQVVCVKFLAINATNNSCACRRSPRGLSPAHLGDGSS 418	PA (EDIN/) EDINGER S R.
QY	301 DLILIRKSRFNFLRLIRHTNQDQ 326	PA (ELLE/) ELLERMAN K.
Db	419 DLILIRKSRFNFLRLIRHTNQDQ 444	PA (MALY/) MALYANKAR U M.
		FA (ORTT/) ORT T.
		FA (GORM/) GORMAN L.
		FA (ZERH/) ZERHUSEN B D.
		FA (ANDE/) ANDERSON D W.
		PA (ZHON/) ZHONG M.
		PA (CATT/) CATTERTON E.
		PA (JIWW/) JI W.
		FA (MILL/) MILLER C E.
		FA (RAST/) RASTELLI L.
		PA (STON/) STONE D J.
		PA (PENA/) PENNA C E A.
		PA (SHEN/) SHENOY S G.
		PA (SHIM/) SHIMKETS R A.
		FA (ROTH/) ROTHENBERG M E.
		FA (LEAC/) LEACH M D.
		PA (AGEE/) AGEE M L.
RESULT 9		
ADN62845		
ID	ADN62845 standard; protein; 537 AA.	
XX		
AC	ADN62845;	
XX		
DT	01-JUL-2004 (first entry)	
XX		
DE	Human NOV9a.	

PA (BERG/) BERGHS C.  
PA (DIP/) DIPPO V A.  
PA (EISE/) EISEN A.  
PA (GANG/) GANGOLLI E A.  
PA (RIEG/) RIEGER D K.  
PA (SPAD/) SPADERNA S K.  
XX  
PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;  
PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;  
XX  
DR WPI; 2004-213931/20.  
DR N-PSDB; ADN62844.  
XX  
PT Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.  
XX  
XX Claim 1; SEQ ID NO 40; 395pp; English.  
XX  
CC The invention relates to isolated NOVX polypeptides and polynucleotides.  
CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
CC treat a medical condition in human related to the aberrant expression and  
CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
CC polynucleotides may be used to treat disorders associated with decreased  
CC expression or activity of NOVX by supplementing the patient our  
CC production or to rectify mutations. Conversely, antisense NA molecules  
CC may be administered to down regulate expression of NOVX polypeptides by  
CC binding with the cells own genes and preventing their expression. NOVX  
CC polynucleotides and complementary sequences may also be used as DNA  
CC probes in diagnostic assays to detect and quantitate the presence of  
CC similar sequences in samples, and so which patients may be in need of  
CC restorative therapy. NOVX polypeptides may also be used as antigens in  
CC the production of antibodies and in assays to identify modulators  
CC (agonists and antagonists) of the expression and activity of NOVX.. The  
CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
CC used to modulate NOVX polynucleotide expression and activity of NOVX  
CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
CC polypeptides and polynucleotides may be used in this way to prevent,  
CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,  
CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
CC disturbances associated with obesity, the metabolic syndrome X and  
CC wasting disorders associated with chronic diseases and various cancers.  
CC They may also be used as antibacterial agents. The present sequence  
CC represents the amino acid sequence of a human NOVX protein.  
XX  
SQ Sequence 537 AA;  
Query Match 95.5%; Score 1640.5; DB 8; Length 537;  
Best Local Similarity 96.3%; Pred. No. 3.5e-164;  
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;  
QY 1 PKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIGNKFYVYVVEVITEHANOAKE 60  
DB 130 PKHLVFINPFGKGQKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 178  
QY 61 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGIIPA 120  
DB 179 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGIIPA 238  
QY 121 GSTDCVCYVTGTSADTSAETSAHIVVGDLSLMDVSSVHHNSTLLRYSVLLGYGFGYDIIK 180  
DB 239 GSTDCVCYVTGTSADTSAETSAHIVVGDLSLMDVSSVHHNSTLLRYSVLLGYGFGYDIIK 298  
QY 181 DSEKKRWLGLARYDPSGLKTFLSHCYEGTVGFLPAQHTVGSFDRKPCRCACFCVCRQSK 240  
DB 299 DSEKKRWLGLARYDPSGLKTFLSHCYEGTVGFLPAQHTVGSFDRKPCRCACFCVCRQSK 358

QY 241 QOLEEEQKALYGLAEAEDEVQVVCGLAINATNMSCACRRSPRGLSPAHLIGDSS 300  
DB 359 QOLEEEQKALYGLAEAEDEVQVVCGLAINATNMSCACRRSPRGLSPAHLIGDSS 418  
QY 301 DLILIRKCSRFNFLRFLIRHTNQDDQ 326  
DB 419 DLILIRKCSRFNFLRFLIRHTNQDDQ 444  
RESULT 10  
ABB07857  
ID ABB07857 standard; protein; 562 AA.  
XX  
XX ABB07857;  
XX  
XX 03-JUL-2002 (first entry)  
XX Human sphingosine kinase-like protein.  
XX  
XX Human sphingosine kinase-like protein; intracellular signalling;  
XX cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;  
XX autoimmune disease; rheumatoid arthritis; Parkinson's disease.  
XX  
XX Homo sapiens.  
XX  
XX WO200228906-A2.  
XX  
XX 11-APR-2002.  
XX  
XX 05-OCT-2001; 2001WO-EP011516.  
XX  
XX 06-OCT-2000; 2000US-0238005P.  
XX 23-AUG-2001; 2001US-0314113P.  
XX  
XX (FARB ) BAYER AG.  
XX  
XX Kossida S, Encinas J;  
XX  
XX WPI; 2002-340094/37.  
XX N-PSDB; ABL40828.  
XX  
XX New reagent for modulating the activity of sphingosine kinase-like  
XX protein polypeptide or polynucleotide and treating cancer, asthma,  
XX allergy, an autoimmune disease, or a central or peripheral nervous system  
XX disorder.  
XX  
XX Claim 25; Fig 11; 120pp; English.  
XX  
XX The invention relates to a human sphingosine kinase-like protein. The  
XX polypeptide can be expressed by standard recombinant methodology. The  
XX sphingosine kinase-like protein and gene can be used to regulate  
XX intracellular signalling and consequently cell proliferation and  
XX apoptosis. Such regulation is useful for treating cancer, allergies (e.g.  
XX asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and  
XX peripheral nervous system disorders (e.g. Parkinson's disease). The  
XX present sequence represents the human sphingosine kinase-like protein  
XX  
SQ Sequence 562 AA;  
Query Match 95.5%; Score 1640.5; DB 5; Length 562;  
Best Local Similarity 96.3%; Pred. No. 3.7e-164;  
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;  
QY 1 PKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIGNKFYVYVVEVITEHANOAKE 60  
DB 155 PKHLVFINPFGKGQKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 203  
QY 61 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGIIPA 120  
DB 204 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGIIPA 263  
QY 121 GSTDCVCYVTGTSADTSAETSAHIVVGDLSLMDVSSVHHNSTLLRYSVLLGYGFGYDIIK 180

```

Db 264 GSTDCVSVTGTSDAETSAALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGFGYFDIHK 323
Qy 181 DSEKRWGLARYDFSGDKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRAGCFVCRQSK 240
Db 324 DSEKRWGLARYDFSGDKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRAGCFVCRQSK 383
Qy 241 QOLEEOKKALYGLEAAEDVEEQVCGKFLAINATNNSCACRRSPRGLSPAHLGDGSS 300
Db 384 QOLEEOKKALYGLEAAEDVEEQVCGKFLAINATNNSCACRRSPRGLSPAHLGDGSS 443
Qy 301 DLILIRKCSRFNFLRLIRHTNQDQ 326
Db 444 DLILIRKCSRFNFLRLIRHTNQDQ 469

RESULT 11
AAW49115
ID AAW49115 standard; protein; 537 AA.
XX AC AAW49115;
XX DT 20-MAY-2002 (first entry)
XX DE Human ceramide kinase hCERK1.
XX KW Human; ceramide kinase; hCERK1; drug screening; gene therapy;
KW neurological disease; inflammation; human immunodeficiency virus;
KW HIV infection; type 2 diabetes; obesity; sepsis; arteriosclerosis;
KW cancer; neuroprotective; antiinflammatory; anti-HIV; antidiabetic;
KW anorectic; antibacterial; antiseptic; antiarteriosclerotic; cytostatic;
KW enzyme.
XX OS Homo sapiens.
XX PN WO200196575-A1.
XX PD 20-DEC-2001.
XX PF 11-JUN-2001; 2001WO-JP004889.
XX PR 14-JUN-2000; 2000JP-00178039.
XX PA (SANY ) SANKYO CO LTD.
XX PI Sugiyura M, Kono K, Kohama T;
XX DR WPI; 2002-179513/23.
XX DR N-PSDB; ABA96945.
XX PT Human ceramide kinase gene and the enzyme encoded by it for screening
XX substances as drugs for neurological, inflammatory and other disorders.
XX PS Claim 1; Page 54-57; 61pp; Japanese.
XX CC This sequence represents a human ceramide kinase designated hCERK1. The
CC invention relates to hCERK1, nucleic acids encoding it, expression
CC vectors and host cells containing hCERK1 nucleic acids, the recombinant
CC production of hCERK1 and antibodies specific for hCERK1. The invention
CC also encompasses methods of isolating hCERK1 from samples, the use of
CC hCERK1 in drug screening, and the use of hCERK1 nucleic acid sequences in
CC gene therapy. hCERK1 mediates the ATP-dependent 1-phosphorylation of
CC ceramides and can be used to screen for therapeutic and preventive agents
CC for a wide range of disorders. Such disorders include neurological
CC disease, inflammation, human immunodeficiency virus (HIV) infection, type
CC 2 diabetes, obesity, sepsis, arteriosclerosis and cancer
XX SQ Sequence 537 AA;
XX Query Match 95.1%; Score 1632.5; DB 5; Length 537;
XX Best Local Similarity 95.7%; Pred. No. 2.5e-163;
XX Matches 312; Conservative 1; Mismatches 1; Indels 11; Gaps 1;

Db 130 PKHLVFINPFGKGQKRIYERKVAFLFTLASITTDII :|||||
Qy 61 TLYEINIDKDYGVVCVGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPA 120
Db 179 TLYEINIDKDYGVVCVGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPA 238
Qy 121 GSTDCVSVTGTSDAETSAALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGFGYFDIHK 180
Db 239 GPTDCVSVTGTSDAETSAALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGFGYFDIHK 298
Qy 181 DSEKRWGLARYDFSGDKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRAGCFVCRQSK 240
Db 299 DSEKRWGLARYDFSGDKTFLSHHCYEGTVSFPLPAQHTVGSPRDRKPCRAGCFVCRQSK 358
Qy 241 QOLEEOKKALYGLEAAEDVEEQVCGKFLAINATNNSCACRRSPRGLSPAHLGDGSS 300
Db 359 QOLEEOKKALYGLEAAEDVEEQVCGKFLAINATNNSCACRRSPRGLSPAHLGDGSS 418
Qy 301 DLILIRKCSRFNFLRLIRHTNQDQ 326
Db 419 DLILIRKCSRFNFLRLIRHTNQDQ 444

RESULT 12
ADP55248
ID ADP55248 standard; protein; 531 AA.
XX AC ADP55248;
XX DT 18-NOV-2004 (first entry)
XX DE Human PRO protein sequence SEQ ID NO:1224.
XX KW human; PRO; immune related disease; inflammatory immune response;
KW immune response stimulation; antiallergic; antianaemic; antiairchritic;
KW antiaesthetic; antidiabetic; antinflammatory; antipsoriatic;
KW antirheumatic; antithyroid; CNS; dermatological; gastrointestinal;
KW haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;
KW nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;
KW virucide; gene therapy.
XX OS Homo sapiens.
XX PN WO2004039956-A2.
XX PD 13-MAY-2004.
XX PF 28-OCT-2003; 2003WO-US034381.
XX PR 29-OCT-2002; 2002US-0422472P.
XX PA (GETH ) GENENTECH INC.
XX PI Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;
XX PI Wood WI, Wu TD;
XX DR WPI; 2004-376182/35.
XX DR N-PSDB; ADP55247.
XX CC New PRO polynucleotides and polypeptides, useful in useful in diagnosing
XX and treating an immune related disease, e.g. systemic lupus
XX erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in
XX stimulating an immune response.
XX PS Claim 1; SEQ ID NO 1224; 3009pp; English.
XX CC The present invention describes an isolated PRO nucleic acid (1). Also
XX described: (1) a vector comprising (1); (2) a host cell comprising the
XX vector of (1); (3) a process for producing a PRO polypeptide; (4) an
XX isolated PRO polypeptide; (5) a chimeric molecule comprising the
XX polypeptide of (4) fused to a heterologous amino acid sequence; (6) an
XX antibody which specifically binds to a polypeptide of (4); (7) a

```



```

OS Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX N-PSDB; AAS77728.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX Claim 20; SEQ ID NO 43900; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
XX amino acid sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX Sequence 746 AA;
XX Query Match 81.6%; Score 1400.5; DB 4; Length 746;
XX Best Local Similarity 80.8%; Pred. No. 1.4e-138;
XX Matches 274; Conservative 4; Mismatches 0; Indels 61; Gaps 2;

QY 49 EVITEHANQAKETLYEINIDKYDGVCGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVL 108
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ KLITEHANQAKETLYEINIDKYDGVCGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVL 357
QY 109 VPSSLRIGIIPA-----GSTDCVCYSTVGTSDAETSAIHIV----- 145
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 358 VPSSLRIGIIPAHHVGPGENAGGLDGDSTDCVCYSTVGTSDAETSAIHIVGCCPEARKPP 417
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 146 -----GDSLAMDVSSVHNSLTLRYSV 167
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 418 ASRTACSGHGQCLGCOLNACLCEASRLQSRMQSGDSLAMDVSSVHNSLTLRYSV 477
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 168 SILGYGFYDIIKQSEKKRWGLARYDFSLGKTLFLSHHCYEGTVSFLPAQHTVGSPRDRK 227
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 478 SILGYGFYDIIKQSEKKRWGLARYDFSGIKTLFLSHHCYEGTVSFLPAQHTVGSPRNRK 537
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 228 PCRACGFCVCRSKQGLEEQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPR 287
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 538 PCRACGFCVCRSKQGLEEQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPR 597
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 288 GLSPAHLGDGSSDLILIRKCSRPNFLRLIRHTNQDQ 326
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 598 GLSPAHLGDGSSDLILIRKCSRPNFLRLIRHTNQDQ 636
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 15
ABG13543
ID ABG13543 standard; protein; 727 AA.
XX
XX ABG13543;
XX
XX 18-FEB-2002 (first entry)
XX Novel human diagnostic protein #13534.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX N-PSDB; AAS77730.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX Claim 20; SEQ ID NO 43902; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
XX amino acid sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX Sequence 727 AA;
XX Query Match 77.1%; Score 1323.5; DB 4; Length 727;
XX Best Local Similarity 58.3%; Pred. No. 2e-130;
XX Matches 275; Conservative 0; Mismatches 0; Indels 197; Gaps 3;

QY 52 TEHANQAKETLYEINIDKYDQ-----
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Db	1	TEHANQAKETLYEINIDKYDGYTAHSRMPCCPRNHS	PAGKAAPVQGVGVBEGQRAFPWEP	60
Qy	73	-----	-----IVCVGGDGMFSEVLHGL	89
Db	61	GPKOMPAKQVCRCLPFLKLEASGLLRSEGR	TCRSPDRLCSSCSIVCVGGDGMFSEVLHGL	120
Qy	90	IGRTORSAGVDONHPRAVLVPSSLRIGIIPA-	-----	120
Db	121	IGRTORSAGVDONHPRAVLVPSSLRIGIIPAG	FANDLAGDEVSVLSGPVSGPEGGMVHT	180
Qy	121	-----	-----GSTDCVCYSTVGTSDAETSALH	142
Db	181	FTLVLTALGCEHRSHPHFMDERRTGEHV	PGENAGGLDSTDCVCYSTVGTSDAETSALH	240
Qy	143	IVV-----	-----GDSLAMDVS	154
Db	241	IVVGCPEARKPPASRHTACSGHGQCLG	QCQLMNACLCEASRLQSRMQSPGDSLAM	300
Qy	155	SVHHNSTLLRYSVSLGFGYGDIIKDSE	KRWLGARYDFGLKTFLSHHCYEGTVSFL	214
Db	301	SVHHNSTLLRYSVSLGFGYGDIIKDSE	KRWLGARYDFGLKTFLSHHCYEGTVSFL	360
Qy	215	PAQHTVGSPRDRKPCRAGCFVCRQSKQ	OLEBEEQKALYGLEAAEDVEEWQVCGKFLAIN	274
Db	361	PAQHTVGSPRDRKPCRAGCFVCRQSKQ	OLEBEEQKALYGLEAAEDVEEWQVCGKFLAIN	420
Qy	275	ATNMSCACRRSPRGLSPAHLGDGSSDL	LILIRKCSRNFNRLFLIRHTNQDDQ	326
Db	421	ATNMSCACRRSPRGLSPAHLGDGSSDL	LILIRKCSRNFNRLFLIRHTNQDDQ	472

Search completed: September 3, 2005, 04:37:38  
Job time : 74.5726 secs



**This Page Blank (uspto)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 04:25:44 ; Search time 104.512 Seconds  
(without alignments)  
2753.633 Million cell updates/sec

Title: US-10-631-958-11

Perfect score: 3025

Sequence: 1 HEANGPAPLGVAPPAMRT.....QLVRLFARGIENPKPDSHS 562

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2888	95.5	537	1 CEK1_HUMAN	Q8tct0 homo sapien
2	2428	80.3	531	1 CEK1_MOUSE	Q8k4q7 mus musculu
3	1904	62.9	409	2 Q62PK5	Q62pk5 mus musculu
4	1811	59.9	339	2 Q6NX59	Q6nx59 homo sapien
5	1713	56.6	572	2 Q6GLV1	Q6glv1 xenopus lae
6	1495.5	49.4	485	2 Q6GMF3	Q6gmf3 xenopus lae
7	801.5	26.5	687	2 Q9VNA6	Q9vna6 drosophila
8	763	25.2	487	2 Q9S5L5	Q9s5l5 drosophila
9	666.5	22.0	410	2 Q7PRA8	Q7pra8 anopheles g
10	648	21.4	608	2 Q6USK2	Q6usk2 arabidopsis
11	620.5	20.5	700	2 Q6HG11	Q6hgh1 oryza sativ
12	594	19.6	532	2 Q6UZF6	Q6uzf6 homo sapien
13	583.5	19.3	533	2 Q9LU45	Q9lu45 arabidopsis
14	567.5	18.8	549	2 Q9T7L1	Q9t7l1 caenorhabdi
15	480	15.9	586	2 Q949C3	Q949c3 oryza sativ
16	428	14.1	485	2 Q8L7L1	Q8l7l1 arabidopsis
17	417.5	13.8	1240	2 Q65419	Q65419 arabidopsis
18	373.5	12.3	480	2 Q8H350	Q8h350 oryza sativ
19	370.5	12.2	763	2 Q9LRB0	Q9lrb0 arabidopsis
20	350	11.6	732	2 Q9FHL3	Q9fhl3 arabidopsis
21	343	11.3	280	2 Q6ZP59	Q6zps9 homo sapien
22	331.5	11.0	446	2 Q84501	Q84501 oryza sativ
23	330.5	10.9	757	2 Q94HY9	Q94hy9 oryza sativ
24	330.5	10.9	757	2 Q7XCS9	Q7xcs9 oryza sativ
25	323	10.7	685	2 Q7QLP4	Q7qlp4 anopheles g
26	314.5	10.4	748	2 Q7XN57	Q7xn57 oryza sativ
27	310.5	10.3	641	2 Q9VY78	Q9vy78 drosophila
28	298.5	9.9	661	2 Q9VZW0	Q9vzw0 drosophila
29	291.5	9.6	458	2 Q14159	Q14159 schizosacch
30	277	9.2	654	1 SPH2_HUMAN	Q9nia7 mus musculu
31	267.5	8.8	616	2 Q6AYB2	Q6ayb2 rattus norv

32	267.5	8.8	517	1 SPH2_MOUSE	Q9nia7 mus musculu
33	264.5	8.7	384	1 SPH1_HUMAN	Q9nyal homo sapien
34	264.5	8.7	384	2 Q96HVB	Q96hvb homo sapien
35	264.5	8.7	398	2 Q96GK1	Q96gk1 homo sapien
36	264.5	8.7	470	2 Q8N632	Q8n632 homo sapien
37	260	8.6	388	2 Q88886	Q88886 mus musculu
38	257.5	8.5	624	2 Q86KF9	Q86kf9 dictyosteli
39	254	8.4	381	2 Q91ZN3	Q91zn3 mus musculu
40	254	8.4	382	2 Q8C115	Q8c115 mus musculu
41	254	8.4	504	2 Q88885	Q88885 mus musculu
42	253.5	8.4	579	2 Q74ZE3	Q74ze3 ashbya goss
43	248.5	8.2	687	2 Q06147	Q06147 saccharomyc
44	245.5	8.1	383	2 Q642F6	Q642f6 rattus norv
45	245.5	8.1	423	2 Q7JM91	Q7jm91 caenorhabdi

#### ALIGNMENTS

RESULT 1  
CEK1\_HUMAN  
ID CEK1\_HUMAN STANDARD; PRT; 537 AA.  
AC Q8TCT0; Q9BYB3; Q9UG5;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Ceramide kinase (EC 2.7.1.138) (Acylsphingosine kinase) (hCERK) (Lipid  
kinase 4) (LK4).  
GN Name=CERK; Synonyms=KIAA1646;  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
RC TISSUE=Leukemia;  
RX MEDLINE=22075121; PubMed=11956206; DOI=10.1074/jbc.M201535200;  
RA Sugiura M., Kono K., Liu H., Shimizugawa T., Minekura H., Spiegel S.,  
RA Kohama T.;  
RT "Ceramide kinase, a novel lipid kinase. Molecular cloning and  
functional characterization."  
RL J. Biol. Chem. 277:23294-23300(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Van Veldhoven P.P.;  
RT "A search for lipid kinases."  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20057165; PubMed=10591208; DOI=10.1038/990031;  
RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,  
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,  
RA Bagguley C., Bailey J., Barlow K.P., Bates K.N., Beasley O.P.,  
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,  
RA Burdill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,  
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,  
RA Conroy D., Corby N.C., Coville G.J., Cox A.V., Davis J., Dawson E.,  
RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,  
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,  
RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,  
RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,  
RA Hunt S.E., Jones M.C., Kersey J., Kimberley A.M., King A.,  
RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,  
RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,  
RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,  
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,  
RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,  
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,  
RA Soderlund C., Spraggon L., Steward C.A., Suleston J.E., Swann R.M.,  
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,  
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,  
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,  
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,

RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,  
 RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,  
 RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,  
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.I.,  
 RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,  
 RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,  
 RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,  
 RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,  
 RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,  
 RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,  
 RA Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,  
 RA Korf I., Bedell J.A., Hillier L.W., Mardis E., Waterston R.,  
 RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saito S.,  
 RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,  
 RA Edelman L., Kim U.J., Shizuya H., Simon M.I., Dumaneki J.P.,  
 RA Peyraud M., Kedra D., Seroussi E., Fransson I., Tapia I., Bruder C.E.,  
 RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,  
 RA Khan A.S., Lane L., Tilahun Y., Wright H.;  
 RT "The DNA sequence of human chromosome 22.";  
 RL Nature 402:489-495(1999).  
 RN [4]  
 RP SEQUENCE OF 57-537 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=21156230; PubMed=11258795;  
 RA Hirotsawa M., Nagase T., Murahashi Y., Kikuno R., Ohara O.;  
 RT "Identification of novel transcribed sequences on human chromosome 22  
 by expressed sequence tag mapping.";  
 RL DNA Res. 8:1-9(2001).  
 CC -!- FUNCTION: Catalyzes specifically the phosphorylation of ceramide  
 to form ceramide 1-phosphate. Acts efficiently on natural and  
 analog ceramides (C6, C8, C16 ceramides, and C8-dihydroceramide),  
 to a lesser extent on C2-ceramide and C6-dihydroceramide, but not  
 on other lipids, such as various sphingosines  
 CC -!- CATALYTIC ACTIVITY: ATP + ceramide = ADP + ceramide 1-phosphate.  
 CC -!- COFACTOR: Calcium and magnesium.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.  
 CC -!- TISSUE SPECIFICITY: High level expression in heart, brain,  
 skeletal muscle, kidney and liver; moderate in peripheral blood  
 leukocytes and thymus; very low in spleen, small intestine,  
 placenta and lung.  
 CC -!- MISCELLANEOUS: Optimal pH is 6.0-7.5.  
 CC -!- SIMILARITY: Contains 1 DAGKC domain.  
 CC -!- CAUTION: Ref.3 sequence differs from that shown due to erroneous  
 gene model prediction. An additional exon may exist between amino  
 acid positions 168 and 169.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; AB079066; BAC01154.1; -;  
 DR EMBL; AJ457828; CAD29884.1; -;  
 DR EMBL; AL096766; CAB62977.1; ALT SEQ.  
 DR EMBL; AL118516; -; NOT ANNOTATED\_CDS.  
 DR EMBL; AB051433; BAB3316.1; -;  
 DR Genbank; HGNC:19256; CERK.  
 DR GO; GO:0000299; C:integral to membrane of membrane fraction; IDA.  
 DR GO; GO:0004685; F:calcium/calmodulin-dependent protein kinase. . .; IDA.  
 DR GO; GO:0001729; F:ceramide kinase activity; IDA.  
 DR GO; GO:0000287; F:magnesium ion binding; IDA.  
 DR GO; GO:0006672; F:ceramide metabolism; TAS.  
 DR InterPro; IPR001206; DAGKC.  
 DR InterPro; IPR001849; PH.  
 DR InterPro; IPR011036; PH-related.  
 DR Pfam; PF00781; DAGK\_cat; 1.  
 DR ProDom; PD005043; DAGKC; 1.  
 DR SMART; SM00046; DAGKC; 1.  
 DR SMART; SM00233; PH; 1.  
 KW Calcium; Kinase; Magnesium; Transferase.

FT DOMAIN 132 278 DAGKC.  
 SQ SEQUENCE 537 AA; 59977 MW; 3DBFC0ED8D679F7F CRC64;  
 Query Match 95.5%; Score 2888; DB 1; Length 537;  
 Best Local Similarity 100.0%; Pred. NO. 1.6e-227; Indels 0; Gaps 0;  
 Matches 537; Conservative 0; Mismatches 0;  
 QY 26 MGATGAAPLQSVLVWVKQORCAVSLPARALLRWRSPGCGAGAGADACSVPVSEIIAV 85  
 DB 1 MGATGAAPLQSVLVWVKQORCAVSLPARALLRWRSPGCGAGAGADACSVPVSEIIAV 60  
 QY 86 EETDVHKGQSGKQWKQXMKPYAFTVHCVKRRHRWKWAQVTPWCPEEQICHWLQTLR 145  
 DB 61 EETDVHKGQSGKQWKQXMKPYAFTVHCVKRRHRWKWAQVTPWCPEEQICHWLQTLR 120  
 QY 146 EMLEKLTSRPKHLLVFNPGGKGQGRKRIYERKVAPLFTLASITTDIIIVTEHANAQKETT 205  
 DB 121 EMLEKLTSRPKHLLVFNPGGKGQGRKRIYERKVAPLFTLASITTDIIIVTEHANAQKETT 180  
 QY 206 YEINIDKYDGI VCGVGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGS 265  
 DB 181 YEINIDKYDGI VCGVGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGS 240  
 QY 266 TDCVCYSTGTSDAETSAHIVVGDLSLAMDVSSVHHNSTLLRYSVSLLGFGYGDIIKDS 325  
 DB 241 TDCVCYSTGTSDAETSAHIVVGDLSLAMDVSSVHHNSTLLRYSVSLLGFGYGDIIKDS 300  
 QY 326 EKKRWGLGARYDFSGLTFLSHHCYEGTVSGFLPAQHTVGSPRDRKPCRCAGFCVCRQSKQ 385  
 DB 301 EKKRWGLGARYDFSGLTFLSHHCYEGTVSGFLPAQHTVGSPRDRKPCRCAGFCVCRQSKQ 360  
 QY 386 LEEQKALYGLEAAEDVEEWQVVCGLFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 445  
 DB 361 LEEQKALYGLEAAEDVEEWQVVCGLFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420  
 QY 446 ILIRKCSRNFPLRLIRHTNQDQDFTTFVEVYRVKXQFTSKMEDESDSLKGGKKRF 505  
 DB 421 ILIRKCSRNFPLRLIRHTNQDQDFTTFVEVYRVKXQFTSKMEDESDSLKGGKKRF 480  
 QY 506 GHICSSHPSCCCTVSNSSWNCDEVLHSPALEVVRHCOLVRLFARGIEENPKPDSSH 562  
 DB 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPALEVVRHCOLVRLFARGIEENPKPDSSH 537  
 RESULT 2  
 CEKI\_MOUSE STANDARD; PRT; 531 AA.  
 ID CEKI\_MOUSE  
 AC Q8K4Q7;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Ceramide kinase (EC 2.7.1.138) (Acylsphingosine kinase) (mCERK).  
 GN Name=Cerk;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=22075121; PubMed=11956206; DOI=10.1074/jbc.M201535200;  
 RA Sugiyama M., Kono K., Liu H., Shimizuawa T., Minekura H., Spiegel S.,  
 RA Kohama T.;  
 RT "Ceramide kinase, a novel lipid kinase. Molecular cloning and  
 functional characterization.";  
 RL J. Biol. Chem. 277:23294-23300(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Heart, and Thymus;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaido I., Osato N., Saito R., Saito K., Nishikawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,

RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Nunata K., Okido T., Pavan W.J., Perte G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wymshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs";  
 RL Nature 420:563-573 (2002).  
 CC -!- FUNCTION: Catalyzes specifically the phosphorylation of ceramide  
 to form ceramides 1-phosphate. Acts efficiently on natural and  
 analog ceramides (C6, C8, C16 ceramides, and C8-dihydroceramide),  
 to a lesser extent on C2-ceramide and C6-dihydroceramide, but not  
 on other lipids, such as various sphingosines (By similarity).  
 CC -!- CATALYTIC ACTIVITY: ATP + ceramide = ADP + ceramide 1-phosphate.  
 CC -!- COFACTOR: Calcium and magnesium (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By  
 similarity).  
 CC -!- TISSUE SPECIFICITY: High level expression in heart, brain and  
 testis; low expression in spleen, liver and lung; not detected in  
 skeletal muscle.  
 CC -!- DEVELOPMENTAL STAGE: Highly expressed at 7 dpc and decreases  
 rapidly thereafter.  
 CC -!- SIMILARITY: Contains 1 DAGKc domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AB079067; BAC01155.1; -;  
 DR EMBL; AK042077; BAC31157.1; -;  
 DR EMBL; AK052269; BAC34908.1; -;  
 DR MGD; MGI:2386052; Cerk.  
 DR GO; GO:0000299; C:integral to membrane of membrane fraction; ISS.  
 DR GO; GO:0004685; F:calcium/calmodulin-dependent protein kinase; ISS.  
 DR GO; GO:0001729; F:ceramide kinase activity; ISS.  
 DR GO; GO:0000287; F:magnesium ion binding; ISS.  
 DR GO; GO:0006672; P:ceramide metabolism; ISS.  
 DR InterPro; IPR001206; DAGKc.  
 DR InterPro; IPR011036; PH-related.  
 DR Pfam; PF00781; DAGK\_cat; 1.  
 DR ProDom; PD005043; DAGKc; 1.  
 KW Calcium; Kinase; Magnesium; Transferase.  
 FT DOMAIN 132 278  
 FT CONFLICT 378 378 V -> M (in Ref. 2).  
 FT CONFLICT 467 467 D -> Y (in Ref. 2; BAC34908).  
 FT CONFLICT 467 467  
 SQ SEQUENCE 531 AA; 59811 MW; 31FECC34C348AA0A CRC64;  
 Query Match 80.3%; Score 2428; DB 1; Length 531;  
 Best Local Similarity 85.1%; Pred. No. 7e-190;  
 Matches 450; Conservative 33; Mismatches 46; Indels 0; Gaps 0;

QY 26 MGATGAAPLOSVLVWVKQORCAVSEPARALLRWRSFPGAGAGADACSVPVSEITAV 85  
 DB 1 MGAMGAAPLESVLVWVKRRCAVSEPARALLRWRSFPGAGADARSVLVSEITAV 60  
 QY 86 EETDVHGKHGSGKWKQKKEKYAFVTHCVKARHRWKWAQVTFWCPEQLCHLWLQTLR 145  
 DB 61 EEKDDCEKASSGRHWKKNENFATVHRVKVRHHRWKWARTVFWSADEQLCHLWLQTLR 120  
 QY 146 EMLEKLSRPHKHLVFINPFGKGKGRYRERKVPAPLFTLASITTDIITVTEHANOAKETL 205  
 DB 121 GLLESLSRPHKHLVFINPFGKGKGRYRERKVPAPLFTLASITTEIIITEHANOAKETL 180  
 QY 206 YEINIDKVGIVCGVCGDMFSEVLHGLIGRTORAGVDQNHPRAVLVPSLSRIGIIPAGS 265  
 DB 181 YEINTSDYDGVTCVCGDMFSEVLHGLVIGRTQQSAGIDPNHPRAVLVSTLRIGIIPAGS 240  
 QY 266 TDCVYSTVGTSDAETSAHHTVWGDLSAMDVSSVHHNSTLLRYSVSLGAGYFGYGDIIKDS 325  
 DB 241 TDCVYSTVGTNDATSAHHTVWGDLSAMDVSSVHHNSTLLRYSVSLGAGYFGYGDIIKDS 300  
 QY 326 EKKRWLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCRSQKQ 385  
 DB 301 EKKRWMLGVYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCRSQKQ 360  
 QY 386 LEEQKALYGLAEAEDEVEHQQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 445  
 DB 361 LEEBEKALYGLAEAEVEHQQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420  
 QY 446 ILIRKCSRFRFLRLRHTNQDDQDFTFVSVVRVKFQFTSKHMEDESDSLKEGKKRF 505  
 DB 421 ILIRKCSRFRFLRLRHTNQDDQDFTFVSVVRVKFQFTSKHMEDESDSLKEGKKRF 480  
 QY 506 GHICSSPSCCTVSNSWNCGEVHLSPAIEVVRHCOLVRLFARGIEE 554  
 DB 481 GKICKDRPSCCTVSNSWNCGEVHLSPAIEVVRHCOLVRLFARGIEE 529  
 RESULT 3  
 Q6ZPK5 PRELIMINARY; PRT; 409 AA.  
 AC Q6ZPK5;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE MKIAA1646 protein (Fragment).  
 GN Name=MKIAA1646;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryonic tail;  
 RX PubMed:14621295;  
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,  
 RA Saka Y., Nagase T., Ohara O., Koga H.;  
 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:  
 RT III. the complete nucleotide sequences of 500 mouse KIAA-homologous  
 RT cDNAs identified by screening of terminal sequences of cDNA clones  
 RT randomly sampled from size-fractionated libraries.";  
 RL DNA Res. 10:167-180(2003).  
 DR EMBL; AK129416; BAC98226.1; -;  
 DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.  
 DR GO; GO:0007205; P:protein kinase C activation; IEA.  
 DR InterPro; IPR001206; DAGKc.  
 DR Pfam; PF00781; DAGK\_cat; 1.  
 DR ProDom; PD005043; DAGKc; 1.  
 DR SMART; SM00046; DAGKc; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 409 AA; 45697 MW; 810E815B37E26808 CRC64;  
 Query Match 62.9%; Score 1904; DB 2; Length 409;  
 Best Local Similarity 87.6%; Pred. No. 3.8e-147;

```
Matches 352; Conservative 26; Mismatches 24; Indels 0; Gaps 0;
QY 153 SRPKHLVFNPFPGKGKGIYERKVAFLPTLASITTDIIIVTEHANOAKETLYEINDK 212
Db 6 SRPKHLVFNPFPGKGKGIYERKVAFLPTLASITTDIIIVTEHANOAKETLYEINTDS 65
QY 213 YDGIVCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPAGSTDCVCYS 272
Db 66 YDGIVCVGGDGMFSEVLHGLIGRTQRSAGIDPNHPRAVLPSTLRIGIIPAGSTDCVCYS 125
QY 273 TVGTSDAETSAIHIVVGDLSAMDVSSVHHNSTLLRYSVSLGYPFGYDIIKDSEKKRWLG 332
Db 126 TVGTNDAETSAHIIIGDSLAIQVSSVHHNSTLLRYSVSLGYPFGYDIIKDSEKKRWLG 185
QY 333 LARYDFSLGKTLFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQLEEEQK 392
Db 186 LVRVDFSLGKTLFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQLEEEQK 245
QY 393 ALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGSGSSDLILIRKCS 452
Db 246 ALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGSGSSDLILIRKCS 305
QY 453 RFNLFRLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLKEGKRGFGHICSSH 512
Db 306 RFNLFRLIRHTNQDQDFTFVEVYRVKFKHFTSKHVEDEDNDSKEQEKQKFGKICKDR 365
QY 513 PSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEE 554
Db 366 PSCCSASRWNCDGEVHSPAIEVRVHCOLVRLFARGIEE 407

RESULT 4
Q6NXX59 PRELIMINARY; PRT; 339 AA.
ID Q6NXX59
AC Q6NXX59;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE CERK protein.
GN Name=CERK;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral Nervous System;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral Nervous System;
RA Strausberg R.;
Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; BC067255; AAH67255.1; -.
DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
DR InterPro; IPR001206; DAGKC.
DR ProDom; PD005043; DAGKC; 1.
SQ SEQUENCE 339 AA; 37780 MW; A4C2ACDF2B6F3D0 CRC64;
Query Match 59.9%; Score 1811; DB 2; Length 339;
Best Local Similarity 100.0%; Pred. No. 1-2e-139; Mismatches 0; Indels 0; Gaps 0;
Matches 339; Conservative 0;
QY 224 MFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPAGSTDCVCYSTVGTSDAETSA 283
Db 1 MFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPAGSTDCVCYSTVGTSDAETSA 60
QY 284 LHIIVGDSLAMDVSSVHHNSTLLRYSVSLGYPFGYDIIKDSEKKRWLGARYDPSGLKT 343
Db 61 LHIIVGDSLAMDVSSVHHNSTLLRYSVSLGYPFGYDIIKDSEKKRWLGARYDPSGLKT 120
QY 344 FLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQLEEEQKALYGLEAAEDV 403
Db 121 FLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQLEEEQKALYGLEAAEDV 180
QY 404 EEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGSGSSDLILIRKCSRFNLFRLIRH 463
Db 181 EEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGSGSSDLILIRKCSRFNLFRLIRH 240
QY 464 TNQDQDFTFVEVYRVKFKQFTSKHMEDESDLKEGKRGFGHICSSHSCCTVSNSS 523
Db 241 TNQDQDFTFVEVYRVKFKQFTSKHMEDESDLKEGKRGFGHICSSHSCCTVSNSS 300
QY 524 WNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 562
Db 301 WNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 339

RESULT 5
Q6GLV1 PRELIMINARY; PRT; 572 AA.
ID Q6GLV1
AC Q6GLV1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE MGC84197 protein.
GN Name=MGC84197;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OX Xenopodinae; Xenopus.
RN NCBI_TaxID=8355;
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral Nervous System;
RA Strausberg R.;
Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
```

DT	05-JUL-2004	(TrEMBLrel. 27, Created)
DT	05-JUL-2004	(TrEMBLrel. 27, Last sequence update)
DT	05-JUL-2004	(TrEMBLrel. 27, Last annotation update)
DE	MG81777	protein.
GN	Names=MG81777;	
OS	Xenopus laevis	(African clawed frog).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Eupidae;	
OC	xenopodinae; Xenopus.	
OX	NCBI_TaxID=8355;	
RI	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUS=Embryo;	
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;	
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,	
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,	
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,	
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,	
RA	Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,	
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,	
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,	
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Villation D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,	
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,	
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Villation D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,	
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,	
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Villation D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,	
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,	
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Villation D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,	
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,	
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Villation D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,	
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,	
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Villation D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,	
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,	
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Villation D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,	
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,	
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Villation D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,	
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,	
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Villation D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,	
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,	
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Villation D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,	
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,	
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Villation D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,	
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,	
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Villation D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	

Qy 156 KHLVFINPFGKGGKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLYEINIDKYDG 215  
 Db 158 KYLLVYINPYGKKRGKQIYETKVAPLFSAGIACADVIIVTEVANHARDNLYEVNLEKYDG 217  
 Qy 216 IVCVGGDGMFSEVLHGLIGRTORSAGUNDQNHPRAVLPSSIRIGIIPAGSTDCVCYSTVG 275  
 Db 218 VVCVGGDGIFFSEVLHGLIGRQKSDVDHNNPNAHLSQCNRIRIGIIPAGSTDCICYATVG 277  
 Qy 276 TSDAETSALHIVGDSLAMDVSSVHNHSTLRLYSVSLGGLGYFYGDIIKDSKKWGLGLAR 335  
 Db 278 INDEPETSALHILGDCQPLDVCVSHNKRTFLKYSVSLGGLGYFYGDVLKGTENRWLGPAP 337  
 Qy 336 YDFSLKTLFLSHHCYEGVTSFLPAQHTVGSPRDRKPCRCAGCFVCRSQKQLEEBQKXALY 395  
 Db 338 YDVSCKTFLTHHCYEGSVSFQPAKWLGSPRDQTPCTSGCVICRQSSKQLEEBQKQTVF 397  
 Qy 396 GLE-AAEDVEWQVQVCGHFLAINATNMSCACRSRGLSPAAHLGDGSSDILIRKCSRFP 454  
 Db 398 GSEHRGKODDWTTIKGRFMAINAVMSMCACPRTPKGLSPAHLADGSAADILVRKCSRL 457  
 Qy 455 NFLRLIRHTNQODQDFTEFVEVYRV 480  
 Db 458 DFLRLIRHTSNKQDFPFVEVYRV 483  
 RESULT 7  
 Q9VNA6 PRELIMINARY; PRT; 687 AA.  
 AC Q9VNA6  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
 DE CG16708-PA (CG16708-pb).  
 CN ORNNames=CG16708.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.B.,  
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brattier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,  
 RA Folsler C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jajuli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattet B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skuseki M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; Smith H.O.,  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,  
 RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
 RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila  
 RT melanogaster euchromatic genome sequence.";  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.;  
 RT "The transposable elements of the Drosophila melanogaster euchromatin:  
 RT a genomic perspective.";  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RG FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RG FlyBase;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE003603; AAF52040.2; -  
 DR FlyBase; FBgn0037315; CG16708.  
 DR GO; GO:0048102; P:autophagic cell death; IEP.  
 DR GO; GO:0035071; P:salivary gland cell death; IEP.  
 DR InterPro; IPR001206; DAGK.  
 DR Pfam; PF00781; DAGK\_cat; 1.  
 DR SMART; SM00046; DAGKc; 1.  
 SQ SEQUENCE 687 AA; 75690 MW; AA081230A939412 CRC64;

Query Match 26.5%; Score 801.5; DB 2; Length 687;  
 Best Local Similarity 31.3%; Pred. No. 1.3e-56;  
 Matches 197; Conservative 97; Mismatches 206; Indels 129; Gaps 16;  
 Qy 36 QSVLWVQQRCAVSL--PARALLRWRSPPGAGACGACACSVFVSIIIAVEETDVHGK 93  
 Db 59 QQLVWERLQTKQSPQNEAKPL-----PPDSAPQGGICSYGFSQSHVLHD--DVVSI 111  
 Qy 94 HQSGKQWQMEKP-----YAFTHVCVRARR-----HRWKWAQ 126  
 Db 112 RSGDTKASSLKPPSPGSSSSGCGDVAQKPTSQVLTINYAMRLSKSQSDCNREWLRR 171  
 Qy 127 VTFWCPCEQLCHLWLTIREMLEKLT---SPKHLVFINPFGKGGKRIYERKVAFLF 183  
 Db 172 LTFNSDPYIVRQMDQLQIRKLSSSPTRMRVRLLVFINPYGGRKAGAQTYERHVFIF 231



QY 184 TLASITTDIIIVTEHANOAKETLYEINIDKYGVGCGDGMSEVLHGLIGRTORSAGVD 243  
Db 232 QLAGVNDATCITTORANQVNDIILSHDLGVYDACCVGSGGTVAEINGLIIFROMRELGLD 291  
QY 244 QNHPRAVLPSSLRIGIIPAGSTDCVCYCTVGTGSDAETSALHIVVGDSDLAMDVSSVHNS 303  
Db 292 EQRPPIPRP-ALPVGVPAGSTDTIAYSMTGTADVRTAAIHVILGQHRGLDVCVSNQ 350  
QY 304 TLLRSVSLGFGYGDIIKDEKKRWLGGLARYDFSLGKTLFSLHCHYEGTVSFL----- 357  
Db 351 SLIRFCASVLSYGLGDVAQAQSENRYRMWGPRIEYSGVKAFLNNRGYDAELRLMBEPDLL 410  
QY 358 ---PAQHTVGSPRD-----RKCRACFCVC-----RQSKQOLEEERKKALYGLE 398  
Db 411 LTTPLEDIPQSPDSVCSLGSVPVCYANCQRCSPASSIQEQRSLFTQEESEKEA----E 466  
QY 399 AAEDVE-----EWOVVCCKFLAINAT-----EWOVVCCKFLAINAT 419  
Db 467 RNQOVETEDSHLAASEAALLRRPRPGNLRPTGSISSNRNLGNDQWKVNRGNFMICGA 526  
QY 420 NMSCACRSRPGLSPPAAHLGDGSSDLILIRKCSRNFRLIRHTNQO-DQDFTFVVEY 478  
Db 527 NITCARSPPNGISRYSHLGDGLDILVKTSLNNVRLNTAGRSGDIRNLPFVEY 586  
QY 479 RVKFKQFTSKHMEDESDLKEGKKRFGHC-----SSHPSCCTVSNSSWNCDE 529  
Db 587 RTREFRFRFSAEEDYSLAGS-----COPITPPEEMTAHSS-----STEFSSWNCDE 635  
QY 530 VLHSPAIEVRVHCQLVRLPARGIEENPKP 558  
Db 636 VTTDLDTMRSHCQLIEVPMRGPHYSYK 664

## RESULT 8

Q95S15  
ID Q95S15 PRELIMINARY; PRT; 487 AA.  
AC Q95S15;  
DT 01-DEC-2001 (TremBLrel. 19, Created)  
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)  
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)  
DE HL01538p.  
GN ORFNames=CG16708;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OC NCBI\_TaxId=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkley;  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,  
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY061001; AAL28549.1; -;  
DR FlyBase; FBgn0037315; CG16708.  
DR GO; GO:0048102; P:autophagic cell death; IEP.  
DR GO; GO:0035071; P:salivary gland cell death; IEP.  
DR InterPro; IPR001206; DAGKC.  
DR Pfam; PF00781; DAGK\_cat; 1.  
DR SMART; SM00046; DAGK; 1.  
SQ SEQUENCE 487 AA; 53711 MW; A4E71EC40354BB07 CRC64;

Query Match 25.2%; Score 763; DB 2; Length 487;  
Best Local Similarity 35.5%; Pred. No. 1.2e-53;  
Matches 170; Conservative 73; Mismatches 146; Indels 90; Gaps 10;

QY 154 RPKHLVFNPPGGGQGRKIYERKVAFLTLASITTDIIIVTEHANOAKETLYEINIDKY 213  
Db 2 RVRLLVFNPNPGRKAGAQTVRHRVPFQLAGVDATCITTORANQVNDIILSHDLGVY 61

QY 214 DGIVCVGDGMSEVLHGLIGRTORSAGVDQNHPRAVLPSSLRIGIIPAGSTDCVCYST 273  
Db 62 DAVCCVGGDGTVAEINGLIIFROMRELGLDEORPPYIPRP-ALPVGVPAGSTDTIAYSM 120  
QY 274 VGTSDAETSALHIVVGDSDLAMDVSSVHNSLRLRYSVSLGFGYGDIIKDEKKRWLG 333  
Db 121 HGTADVRTAAIHVILGQHRGLDVCVSNQSLRRCASVLSYGLGDVAQAQSENRYRMWGP 180  
QY 334 ARYDFSLGKTLFSLHCHYEGTVSFL-----PAQHTVGSPRD-----RKPCRACG 376  
Db 181 RRYEYSGVKAFLNNRGYDAELRLMBEPDLLLTTPLEDIPQSPDSVCSLGSVPVCYANC 240  
QY 377 FVC-----RQSKQOLEEERKKALYGLEAAEDVE----- 404  
Db 241 QRCSPASSIQEQRSLFTQEESEKEA-----ERNQOVETEDSHLAASEAALLRRPRPGNLR 296  
QY 405 -----EWOVVCCKFLAINATNMSCACRSRPGLSPPAAHLGDGSSDLILIR 449  
Db 297 LPTGSISSNRNLGNDQWKVNRGNFMICANITCARSPPNGISRYSHLGDGLDILV 356  
QY 450 KCSRNFRLIRHTNQO-DQDFTFVVEYRVKFKQFTSKHMEDESDLKEGKKRFGHI 508  
Db 357 KTSLLNNVRLNTAGRSGDIRNLPFVEYRTREFRFRFSAEEDYSLAGS----- 408  
QY 509 C-----SSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLPARGIEENPKP 558  
Db 409 COPITPPEEMTAHSS-----STEFSSWNCDEVVTTDLDTMRSHCQLIEVPMRGPHYSYK 464

## RESULT 9

Q7PRA8  
ID Q7PRA8 PRELIMINARY; PRT; 410 AA.  
AC Q7PRA8;  
DT 01-MAR-2004 (TremBLrel. 26, Created)  
DT 01-MAR-2004 (TremBLrel. 26, Last sequence update)  
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)  
DE ENSANGP0000010845 (Fragment).  
GN Name=ENSANGG0000008356;  
OS Anopheles gambiae str. PEST.  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.  
OC NCBI\_TaxId=180454;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PEST;  
RA Anopheles Genome Sequencing Consortium;  
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
CC -!- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; AAA801008859; EAA07556.2; -;  
DR GO; GO:0004143; P:diacylglycerol kinase activity; IEA.  
DR GO; GO:0007205; P:protein kinase C activation; IEA.  
DR InterPro; IPR001206; DAGKC.  
DR Pfam; PF00781; DAGK\_cat; 1.  
DR ProDom; PD005043; DAGK; 1.  
FT NON\_TER 1  
FT NON\_TER 410  
SQ SEQUENCE 410 AA; 45669 MW; 79535B4C0D150154 CRC64;

Query Match 22.0%; Score 666.5; DB 2; Length 410;  
Best Local Similarity 38.9%; Pred. No. 7.4e-46;  
Matches 144; Conservative 59; Mismatches 140; Indels 27; Gaps 5;

QY 118 RHRWKMAQVTFWCPEEOLCHLWLQTLREMLEKLTSPKHLVFNPPGGGQGRKIYER 177  
Db 41 KTNRVRVAVALHNSPRLVELWYNRLSSDLRD-QNRPKHLFLFNPFGGKQNALALYR 99  
QY 178 KVAFLTLASITTDIIIVTEHANOAKE--TLYEINIDKYDGVGVGDMGFSEVLHGLIGR 235  
Db 100 YAKPLFRLAGVDINLIITQRAQIYDVTSKSILLDNDYDLGVCCGSDGTFAELFNLVTR 159

```

Qy 236 TORSAGVDQNHPRAVLVPSLRIGIIPAGSTDCVCYSTVGTSDAETSALHIVVGDSLAMD 295
Db 160 TMDGDIDIKYP-AVLPRNPPIGIVIPAGSTDTVACCLNGTTDIKTCTIHIILGOHSLD 218
Qy 296 VSSVHH-----NSTLLRYSVSLGCGFYGDIIKDSEKKRWGLGARY 336
Db 219 ISAVYSADAAAKCDGASPGAGTGRPRQLLKFASALSYGLGDIAYDSEKYRWMPKRY 278
Qy 337 DFSGLKTFSHHCYEGTVSFPLPAQHTVGSPPDRKPCRAGCFVCRSQKQLEEBOKKALYG 396
Db 279 DYSGFKFLAWGYNAEIIVHLDRRGQDPNDGVRCCLKARCAKAKYGRDCGGERASY- 337
Qy 397 LEAEDVEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSGLILIRKCSRNF 456
Db 338 ---EDDDTEPLVVRGKFLMVSGANISCSERSPOGSPYCHLGDGLDLVLVRHTSMFN 394
Qy 457 LRFLIRHTNQ 466
Db 395 LRLLLTWTSTK 404

RESULT 10
Q6USK2 PRELIMINARY; PRT; 608 AA.
AC Q6USK2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Ceramide kinase.
GN Name=CBRK;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_taxid=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22959668; PubMed=14563678; DOI=10.1101/gad.1140503;
RA Liang H., Yao N., Song J.T., Luo S., Lu H., Greenberg J.T.;
RT "Ceramides modulate programmed cell death in plants.";
RL Genes Dev. 17:2636-2641(2003).
DR EMBL; AY362552; AAQ62904.1; -.
DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
DR InterPro; IPR001205; P:protein kinase C activation; IEA.
DR Pfam; PF00781; DAGK_cat; 1.
DR ProDom; PD005043; DAGKC; 1.
KW Kinase.
SQ SEQUENCE 608 AA; 68462 MW; 65BB9353692D9CA6 CRC64;

Query Match 21.4%; Score 648; DB 2; Length 608;
Best Local Similarity 31.6%; Pred. No. 4.2e-44;
Matches 182; Conservative 68; Mismatches 188; Indels 138; Gaps 19;

Qy 77 VPVSEIIAVEETDV-----HGKHQSGKQKMEKPYAFTVHCVRARRH--RWK 123
Db 71 IKFSDIYAVEFVSYGLVHSPKGLRHAKCEFRERLLNTQEMRYFTVHGFSQSPKPCLDWN 130
Qy 124 WAQVTFWCPBEQCHLWLOTREMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLF 183
Db 131 LAAFTFGHMDLQTCQSWMDQLNYSYLIKEVERPRNLLVFVHPKSGKNGSKVME-TVSKIF 189
Qy 184 TLASITTDIIIVTEHANOAKETLYEI---NIDKYDGI VCVGDMGFSEVLHG-LIGRTO-- 237
Db 190 ITRAKNTKVIIVTERAGHAFDWNASIQNKELHTYDGLIIVAGDGFNEILNGVILLSRLKVP 249
Qy 238 -----RSAG-----VDQ--NHPRAVLVPSL----- 256
Db 250 LPPSPDSFNSVQSRGSSVPEPGDVHETDQKEHY-LLPDSVQVEMNFRVTNGSCE 306
Qy 257 -----RIGIIPAGSTDCVCYSTVGTSDAETSALHIVVGDSLAMDVSV--- 299
Db 307 GIEDPDHPFSRPRFGLIPAGSTDAI VNCITTGARDPVTSAHLIILGRKFLDAMQVVRW 366

```

```

Qy 300 --HHNSTL---LRYSVLLGCGFYGDIIKDSEKKRWGLGARYDFSLKTLFSLHHCYEGTV 354
Db 367 KTASTSTIEPIRYAASFAGVGYGDVI SESEKRYWMPKRYDVGTGKIFLKHRSYEAEV 426
Qy 355 SFPLAQ-----HT-----VGSPPDRK-PCRAGCFVCRSQKQLEEBOKKALYLE 398
Db 427 MFEAESENSKASLHTRSKTWPFRNTTTRSEKILCRANCKICNS-----KVGWNSASTTLN 481
Qy 399 AAEDEVWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSGLILIRKCSRNFRLR 458
Db 482 PCPEKTKWRTKGRFLSIGAAVMSNRNERAPDGLVVD AHLSDGFLHULLIKDCSRPKYLW 541
Qy 459 FLIRHTNQ-ODQDFTFEVYVRVKFQFTSKHMEDESDLKEGKGRFGHICSHSPSCCC 517
Db 542 HLTELAKRGGEPLNFVEYHKTRAFFTS-----FG----- 573
Qy 518 TVSNSWNCDEVLHSPAIEVRVHVCQLVRLFARGIE 553
Db 574 --EESVWNLDEIPEAHQLSAQVLRGLIPLFASGPE 607

```

## RESULT 11

```

Q6H6H1 PRELIMINARY; PRT; 700 AA.
AC Q6H6H1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative ceramide kinase.
GN Name=P0519E06.23; Synonyms=OJ1003.B06.39;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_taxid=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, PAC clone: P0519E06.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005006; BAD25678.1; -.
DR EMBL; AP004676; BAD25337.1; -.
DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
DR InterPro; IPR001206; DAGKC.
DR ProDom; PD005043; DAGKC; 1.
KW Kinase.
SQ SEQUENCE 700 AA; 77961 MW; FBB2C2A3DF31C75D CRC64;

Query Match 20.5%; Score 620.5; DB 2; Length 700;
Best Local Similarity 28.1%; Pred. No. 9e-42;
Matches 193; Conservative 87; Mismatches 197; Indels 209; Gaps 26;

Qy 18 WRTSPAAB---MGATGAAPLOSVLVWVQKQCAVSLP-ARALLRWFRSPFGAGAPCAD 73
Db 59 WLMKSSGDRYLGQHDIEEVSSCW-----SSIMQPKLESKLKF-----SD 101
Qy 74 ACSVPVSRIIAVEE-----TDVHGKQSGKQKMEKPYAFTVHCVRARRHVKW--AQ 126
Db 102 VYAVELLEVPVCEPFWNARATVQKIN-----TENR---FVIHTVTRKRPSPVPC 153
Qy 127 VTFWCPBEQCHLWLOTREMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFLA 186
Db 154 YIFGHKQDQCTKTWVEHIKTCINKEQDRPKSLMVFVHPLCGKGRCKNWE-TVAPLPERA 212
Qy 187 SITTDIIIVTEHANOAKETLYEI---NIDKYDGI VCV----- 219

```

Db 213 KVKTKVIVTQRAGHAVDTLASLSDKDLKKFGDVIAVNTINACLISLFDIKHNYKMSARPE 272  
QY 220 -----GGDGMFSEVL----- 229  
Db 273 NTLSDYDQSAAGSHKSMILFYCFIINNMMKQBHRNNDLSNSELTGDDANAISGSSNTPDD 332  
QY 230 HGLIGRTORSAGVD-----QNHPRVLV--PSS-LRIGIIPAGSTDC-----V 269  
Db 333 HEPLISTTRSTGLDIISSDSSDEPCNGDQVPLVSPFNWFRUGIIPSGSTDAIVLSPVDV 392  
QY 270 CYSTVGTSDAETSALHIVVGDSLAMDVSSV-----HNNSTLLRYSVSLIGYFGYGD1 321  
Db 393 C-STTGERDPVTSALLIILGRISLSDIAQVVRWKSPSAEVSPTVRYAASFAGYFGYGEV 451  
QY 322 IKDSEKKRWLGARVDFSGLTKFLSHHCYEGTVSFLPAQHT-----V 363  
Db 452 IRESEKRYMGPARYDFSGTWFLKHSRYEAKVAFLNGNTHSLTASAEENNANGVQTLQY 511  
QY 364 GSPDRPK-CRAGCFVCR---QSKQLEEEQKALYGLEAAEDVEWQVCGKFLAINAT 419  
Db 512 HONRHKTKICRNTCLIKGTSTSEQSEDENPDS---SRTACETPKVWMSKGRFLUSVGA 568  
QY 420 NNSCACRSPRGLSPAHLGDSSDLILIRKCSRNFNRLRHTNQ-----QDQDFDTF 474  
Db 569 VLSNERAPDGLVADAHLSDFLHLILRDCP-----LFFYLWHLTQTKKGSDFLSKPF 624  
QY 475 VEVYRVKQFQTKMGWEDSDLDKGGKKRFGHICSSHPSCCCTVSNSSWNCDEGVLHSP 534  
Db 625 VEHHTQTAFTFISSHDE-----SVNNLDGELLQAC 654  
QY 535 AIEVRVHCQLVPLFARGIE--ENPKP 558  
Db 655 EVSVOAQRGLVNLFSAGPEKMENTAP 680

RESULT 12  
OSUZP6 PRELIMINARY; PRT; 532 AA.  
ID Q6UZP6  
AC Q6UZP6; TREMBLrel. 27, Created)  
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
DE Ceramide kinase-like protein.  
GN Name=CERK;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=14691825;  
RA Tuson M., Marfany G., Gonzalez-Duarte R.;  
RT "Mutation of CERKL, a novel human ceramide kinase gene, causes  
RT autosomal recessive retinitis pigmentosa (RP26).";  
RL Am. J. Hum. Genet. 74:128-138 (2004).  
DR EMBL; AY357073; AAR13670.1; -.  
DR GO; GO:0004143; P:diacylglycerol kinase activity; IEA.  
DR GO; GO:0007205; P:protein kinase C activation; IEA.  
DR InterPro; IPR001206; DAGKC.  
DR Pfam; PF00781; DAGK\_cat; 1.  
DR ProDom; PD005043; DAGKC; 1.  
KW Kinase.  
SQ SEQUENCE 532 AA; 59602 MW; C73E590F7C25BED1 CRC64;  
Query Match 19.6%; Score 594; DB 2; Length 532;  
Best Local Similarity 28.4%; Pred. No. 9.1e-40;  
Matches 159; Conservative 107; Mismatches 226; Indels 68; Gaps 14;  
QY 9 PLGVAPPAPAWRTSPAEMGATGAAP--LQSVLWVKQRCVSLSEFAPARALLRWMSRPGF 66  
Db 23 PEAAVPPALLTSP---QQTAAAEAILLRGIFEGIRGSCDVLS-ERALR--WRPIQPE 76  
QY 67 AGAGADACSVPSVEIIAVEET-DVHGKHQGGKQKMEKPYAFTVHCVKRARRHWKWA 125

Db 77 RPAGSKYDLCKEETFEIKDIFSVLKRKRSVQQRSGTLLGITLFTCLKKQBKNLQNS 136  
QY 126 QVTWCPPEQLCHLWQLTREMLEKLTSRPKHLVFINPFGGKGQGRKRIYERKVAPLFTL 185  
Db 137 TLDLINLSEDCDIWFRQFKILAGFPNRPKSLKILLNPQSHKKEATQVYVEKVEPLKL 196  
QY 186 ASIITDIIIVTHANQAKETLYEINIDKYDGVICVCGDGMFSEVLHGLIGRTORSAGVDON 245  
Db 197 AGIKTDVTIMYEGHALSLKCELOQFGDGVVCGDGSASEVAHALLRQAQKNAGMETD 256  
QY 246 HPRVLVP--SSLRIGIIPAGSTDCVCTVSTVGTSDAETSALHIVVGDLSLAMDVSSVHNS 303  
Db 257 R---ILTPVRAQLPLGLIPAGSTVNLASHLGVPHVITATLHIINGHVLQVDVCTFSTAG 313  
QY 304 TLLRYSVLLGYGFYGDIIKDSKKRWLG-LARYDFSGLTKFLSHHCYEGTVSFLPAQHT 362  
Db 314 KLLRFGFSAM-FGFGRTLALAEKYRWMSPNQRDFAVVKALAKLAEDCEISFLPFNSS 372  
QY 363 VGSFDRPKRAGCFVCRQSKQLEEEQKALYGLEAAEDVEWQVCGKFLAINATNWS 422  
Db 373 -----DDVOERRAQGSPKSDCNDQMOMIQOQFLNVSMAIP 408  
QY 423 CACRRSPRGLSPAHLGDSSDLILIRKCSRNFNRLRHTNQDQDQDFTFVEYRVKX 482  
Db 409 CLCSVAPRGLAPNTRLNNGSMALTIARNTSRPEFIKHLKRYASVKNQNFPPFVETYTEE 468  
QY 483 FQFTSKH-----MEDESDLKEGGKKRFGHICSSHPSCCCTVSNSSWNCDEGVLH-SP 534  
Db 469 VKVHPNNTGGVPEEEDETASE-----NC-----PPNVVDGLMEVAS 508  
QY 535 AIEVRVHCQLVPLFARGIEE 554  
Db 509 EVHRLHPRLISLYGGSME 528

RESULT 13  
Q9LU45 PRELIMINARY; PRT; 533 AA.  
ID Q9LU45  
AC Q9LU45; (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
DE Similarity to unknown protein.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20181125; PubMed=10718197;  
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,  
RA Tabata S.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence  
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC  
RT clones.";  
RL DNA Res. 7:31-63 (2000).  
DR EMBL; AB023044; BAA97392.1; -.  
DR GO; GO:0004143; P:diacylglycerol kinase activity; IEA.  
DR GO; GO:0007205; P:protein kinase C activation; IEA.  
DR InterPro; IPR001206; DAGKC.  
DR ProDom; PD005043; DAGKC; 1.  
SQ SEQUENCE 533 AA; 60153 MW; 1B851C7606B03E0E CRC64;  
Query Match 19.3%; Score 593.5; DB 2; Length 533;  
Best Local Similarity 33.5%; Pred. No. 6.6e-39;  
Matches 158; Conservative 57; Mismatches 157; Indels 99; Gaps 16;  
QY 77 VPVSEIIAVERTDV-----HGKHQGGKQKMEKPYAFTVHCVKRARRH--RWK 123  
Db 71 IKFSDIYAVEFVSVGLVHSPKGLGRHAKCEPRLNLTQEMRYFTVHGFSQSPKEPCLWN 130

```
Qy 124 WAQVTFWCPBEQLCHLWLTQLEMLEKLTSRPKHLVFINPFGKGQGRKIYERKVAFLF 183
Db 131 LAAPTFGHWDIQTCSWMDQLNYSLLIKEVERPNLLVFHPKSGKNGSKWE-TVSKIF 189
Qy 184 TLASITTTIIVTEHANOAKETLYEI---NIDKYGIVCVGGGMPSEVLHG-LIGRTQ-- 237
Db 190 IRAKVNTKVIIVTERAGHAFDWMASIQNKELHTYDGIITAVGGDGFNETINGLLSLRKVP 249
Qy 238 -----RSAG-----VDQ--NHPRAVLVPSL----- 256
Db 250 LPPSPDSFNSVQSGSSVPEPDGEVHETDQKEHP---LLPDSVQVMNFRIEDPDHP 306
Qy 257 -----RIGIIPAGSTDCVYSTVGTSDAETSALHIVVGDSLAMDVSSV-----HNNSTL 305
Db 307 FSSERPRGLIPAGSTDAIVMCTTGARDPVTSAHLIILGRKFLDAMQVVRWKTASTITI 366
Qy 306 ---LRYSVSLGIGYGYGDIIDKSEKKRWGLARYDFSGLKTFLSHHCHYEGTVSFLPAQ-- 360
Db 367 EPIRYAASAFAGYGYGDIIVSESEKRYWVGPRYDVYGTGKIFLKHRSYEAEMFEEASE 426
Qy 361 -----HT-----VGSPRDK-PCRAGCFVCRSQKQOLEEOKKALYGLEAAEDVEEW 406
Db 427 NSKASLHRTKSTWPPRNTTRSEKILCRANCKICNS-----KVGWNSATTLNCPCKETRW 481
Qy 407 QVVCCKFLAINATNMNSCACRRSPRGLSPAHLGSGSDLIILIRKCSRNFNL 457
Db 482 CRTKGRFLSIGAVNSNENRAPDGLVVDVADHLSDGFLHLIILKDCSRPKYL 532
RESULT 14
Q9TZI1
ID Q9TZI1 PRELIMINARY; PRT; 549 AA.
AC Q9TZI1.
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein T10B11.2.
GN Name=T10B11.2; ORFNames=T10B11.2;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=93069613; PubMed=9851916;
RG WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Minx P., Kemp K.;
RT "The sequence of C. elegans cosmid T10B11.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF098993; AAC67466.1; -.
DR PIR; T33517; T33517.
```

```
DR WormBase; WBGene00020398; T10B11.2.
DR WormPep; T10B11.2; CE18241.
DR InterPro; IPR001206; DAGKC.
DR Pfam; PF00781; DAGK_cat; 1.
KW Hypothetical protein.
SQ SEQUENCE 549 AA; 62425 MW; DE9573755534EBB CRC64;

Query Match 18.8%; Score 567.5; DB 2; Length 549;
Best Local Similarity 30.8%; Pred. No. 1.4e-37;
Matches 139; Conservative 91; Mismatches 160; Indels 61; Gaps 13;

Qy 114 VKRARRHRKWAQ--VTFWCPBEQLCHLWLTQLEMLEKLTSRPKHLVFINPFGKGQ 171
Db 123 VYKDKQKWLKQIPVIFYTTSR--DYWHSLLDTTLRRVKNRKNIIIFINPFGNGKA 180
Qy 172 KRIYERKVAFLFLA-SITTDIIVTEHANOAKETLYEINIDKY---DGIYCVGGGMPSE 227
Db 181 QKIFKDNVDAPFWLTPLGLRYKVLTERANHARDYIVEMPPQWSAIDGLVSGDGLFNE 240
Qy 228 VLHGLIGRTORSAGVDQNHPRAVLVPSLSRIGIIPAGSTDCVYSTVGTSDAETSALHIV 287
Db 241 LLSGALLRTQTDAGRINDPSSHLVTPHIRFGIIGAGSANSIVSTVHETNDHATSAVHIA 300
Qy 288 VGDSLAMDVSSVHNNSTLLRYSVSLGIGYGYGDIIDKSEKKRWGLARYDFSGLKTFLSH 347
Db 301 IGSECNVDVCTVHQHQLIRISANAISYGLWGLDVRDSEYRCLGPIRYQWSALRTTIRH 360
Qy 348 HCYEGTVSFLPAQHTVGSPPDR-KPCRAGCFVCRSQKQOLEEOKKALYGLEAAEDVEEW 406
Db 361 PIYRGMVQFSLSHKENYNPKDLPCLPEPCVCMK-----PQNDKDYDWHHA-----EFT 411
Qy 407 QVVCCKFLAINATNMNSCACRRSPRGLSPAHLGSGSDLIILIRKCSRNFNLRLIRHT-- 464
Db 412 HVICCVPTVTPP-----TPYGLAPFTGIGDTLIDLALVPRISFHNQFMRKVAMY 463
Qy 465 --NQDQDFDTFVEVYRVKKFQFTSKHMEDESDLKEGGKKRFGHICSSHPSCCCTVSNS 522
Db 464 GKGQLYELD-PSLNCYRVTKWSY-----QPDADQEDRG----- 495
Qy 523 SWNCDGCVLHSP---AIEVRVHCOLVRLFAR 550
Db 496 VWNLDGEITLQPKDEPLHLFKLHPQLISFFGR 526

RESULT 15
Q949C3
ID Q949C3 PRELIMINARY; PRT; 586 AA.
AC Q949C3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein C875ERIPDS.
GN Name=C875ERIPDS;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21329048; PubMed=11435398; DOI=10.1101/gr-1617R;
RA Mayer K., Murphy G., Tarchini R., Wambutt R., Volckaert G., Pohl T.,
RA Dueterhoeft A., Stiekema W., Entian K.D., Terryn N., Lemcke K.,
RA Haase D., Hall C.R., van Dodeweerd A.M., Tingey S.V., Mewes H.W.,
RA Bevan M., Bancroft I.;
RT "Conservation of microstructure between a sequenced region of the
genome of rice and multiple segments of the genome of Arabidopsis
thaliana.";
RL Genome Res. 11:1167-1174(2001).
DR EMBL; AJ307662; CAC39069.1; -.
DR Gramene; Q949C3; -.
DR GO; GO:0004143; P:diacylglycerol kinase activity; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
```



**This Page Blank (uspto)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 6, 2005, 07:01:07 ; Search time 5404.27 Seconds  
(without alignments)  
3958.373 Million cell updates/sec

Title: US-10-631-958-11

Perfect score: 3025

Sequence: 1 HEANGPAPLGVAPPAWRT.....QLVRLFARGIEENPKPDHS 562

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US10631958/runat.02092005.165816.4562/app.query.fasta\_1.1941  
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40 cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10631958 @CNC 1.1 9235 @runat.02092005.165816.4562 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_htc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsa1:\*  
9: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2478.5	81.9	1824	3 AK042077	Mus muscu
2	2469.5	81.6	4248	3 AK052269	Mus muscu
3	1518	50.2	1063	4 BM479389	AGENCOURT
4	1367	45.2	1059	5 BQ054406	AGENCOURT
5	1321	43.7	1078	5 BQ063738	AGENCOURT
6	1299.5	43.0	797	7 CK000755	AGENCOURT
7	1284	42.4	1047	5 BQ057191	AGENCOURT
8	1275	42.1	713	7 CK000525	AGENCOURT
9	1260	41.7	820	6 CD655311	AGENCOURT

10	1244	41.1	732	7	CF135528
11	1205	39.8	758	6	CB246749
12	1181	39.0	653	5	BX952302
13	1170	38.7	661	7	CN296312
14	1107	36.6	1017	5	BQ879739
15	1059	35.0	812	7	CK603033
16	1040	34.4	582	5	BP274786
17	1012	33.5	581	5	BP224560
18	1008	33.3	550	2	AW499858
19	997	33.0	584	5	BP310011
20	986	32.6	779	5	BQ265377
21	982	32.5	584	5	BP309990
22	971	32.1	763	7	CK364327
23	968	32.0	573	7	CF138275
24	964	31.9	819	7	CN235868
25	959	31.7	550	7	CF138634
26	943.5	31.2	759	5	BU264162
27	942	31.1	541	7	CN296311
28	925	30.6	725	5	BU214294
29	924	30.5	607	6	CB581157
30	916	30.3	576	6	CB608292
31	914	30.2	653	6	BY754042
32	909	30.0	608	6	CB580936
33	876	29.0	553	6	CA578972
34	874	28.9	662	2	BB638867
35	873	28.9	674	1	AJ739034
36	849	28.1	902	5	BU373554
37	841	27.8	656	6	BY739924
38	838	27.7	545	6	CB611947
39	828	27.4	488	5	BU430459
40	813	26.9	540	6	CB612897
41	811	26.8	474	2	AW503999
42	806	26.6	638	1	AJ739022
43	805	26.6	449	2	AW503483
44	797.5	26.4	759	5	BU290807
45	786.5	26.0	680	2	BB235873

#### ALIGNMENTS

AK042077 1824 bp mRNA linear HTC 03-APR-2004  
Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched library, clone: A630056D11 product: DA59H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS) (FRAGMENT) homolog (Homo sapiens), full insert sequence.

AK042077 AK042077.1 GI:26334912

VERSION HTC; CAP trapper.

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE

AUTHORS

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

PUBMED

REFERENCE

AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,



Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)  
5  
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 1824)

## JOURNAL

## REFERENCE

## AUTHORS

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Niehi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.  
Direct Submission

## TITLE

## JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

## COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.jp/  
URL: http://fantom.gsc.riken.jp/.

## FEATURES

## source

Location/Qualifiers  
1..1824  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="PANTOM\_DB:A630056D11"  
/db\_xref="taxon:10090"  
/clone="A630056D11"  
/tissue\_type="thymus"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="3 days neonate"  
85..1680

## CDS

/note="unnamed protein product; DA59H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS) (FRAGMENT) homolog [Homo sapiens] (SPTP[Q9UE5, evidence: FASTV, 86.5%ID, 100%length, match=945) putative"  
/codon\_start=1  
/protein\_id="BAC31157.1"  
/db\_xref="GI:26334913"  
/translation="MGAMGAAEPLHSLVWVKRRRCVSLSEPARALLRWRRSPFPGPSA

PGDARSVLVSEIIAIVEKDDCKHASSGRWHKVENPFAFVHRVKKVRHHRWKARV  
TFWSADEQLCHLWQLRLGLESLTSPKHLVFINPFGKGQCKRIYKRTVAPLFTL  
ASITTEIITTEHANOAKETIYENTSDYDGVICVGGDMFSEVLHGVIKRTVQSGAGID  
PNHPRVLVPSTLRIGIIPAGSTDCVYCTVGTNDAPTSALHIIIGDSLADIVSSVHV  
HNTLLRYSVLLGFGYGDLLKDKSEKRWMLVRYDFSGUKTFLSHQIYEGTSLFLPA  
QHTVGSFRDNKPCRAGFCVCRQKQLEEEKKALYGLNAEEMEEQVTCGKGFLLAN  
ATNMSCACPRSPGSLFPFAHLGSGSLLIILIRKSRFNFLRLIHRHNTQDQDFTFV  
EVYRVKKFHTSKHVEDNDNSKEQKQKFGKICKDRPSCSASRSSMNCDEGVMS  
PAIEVRVHCQLVRLFARGIEES"

## ORIGIN

Alignment Scores:  
Pred. No.: 1,09e-219 Length: 1824  
Score: 2478.50 Matches: 470  
Percent Similarity: 89.05% Conservative: 34  
Best Local Similarity: 83.04% Mismatches: 54  
Query Match: 81.93% Indels: 8  
DB: Gaps: 1  
US-10-631-958-11 (1-562) x AK042077 (1-1824)  
Qy 3 AlaAlaAsnGlyProAla-ProLeuGlyValArgAlaProProAlaTrpArgThrSerPr 22  
Db 13 GCCGCTAACGGTCCCGCGCCCCCGGTGTCGCGCGCCCGCGCGCGCGCGAGGCC 72  
Qy 22 o--AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpVal 41  
Db 73 GCTCTTGGCGCATGGGGGCAATGGGGCGGGGCGCGTGCATCCGTCTGGGTG 132  
Qy 42 LysGlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArg 61  
Db 133 AAACGGCGCGCGTGTGCGGTGAGCGCGCGCGCGCGTGTGTCACGTGCGTGGCGG 192  
Qy 62 SerProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGlu 81  
Db 193 AGCCCGAGGCGCGCGCGCTCCGCGCGGGTGCCTGCGGTACTGCTGGTGTGGAG 252  
Qy 82 IleIleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGln 101  
Db 253 ATCATCGCTGTTGAGGAAAGACGACGCGAGAAACACGCGCTCAGTGGCGCATGCAT 312  
Qy 102 LysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgHisArg 121  
Db 313 AAAATGGAATATCGTTCCGATTCCAGTCCACGTCACCGTGAAGCGAGTACGACATCAGCGC 372  
Qy 122 TrpLysTrpAlaGlnValPheTrpCysProGluGlnGlnLeuCysHisLeuTrpLeu 141  
Db 373 TGGAAAGTGGGACGGGTGACCTTCTGGAGCGCGCGAGCAGCGTGTGTACCTGTGGCTG 432  
Qy 142 GlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPhe 161  
Db 433 CAGACCTTCGTGGGTGCTGGAGAGCTGACTTCAGAGCCGAGACGACTTGTGTATTTC 492  
Qy 162 IleAsnProPheGlyLysGlyGlnGlyLysArgIleTrpGluArgLysValAlaPro 181  
Db 493 ATCAACCTTCGGAGGAAAGGTGAGGCAAGCGCATCTATGAAAAAACAGTGGCGCT 552  
Qy 182 LeuPheThrLeuAlaSerIleThrThrAspIleValThrGluHisAlaAsnGlnAla 201  
Db 553 CTGTTTACCTTGGCTTCCATCCTACGAGATCATCAATACAGAGCATGCCAACAGGCC 612  
Qy 202 LysGluThrLeuTrpGluIleAsnIleAspLysTyrAspGlyIleValCysValGly 221  
Db 613 AAGGAGACTTATACGAGATCAACACAGACGACTATGATGGCATCGTGTGGTAGGTGG 672  
Qy 222 AspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGly 241  
Db 673 GACGGCATGTTACGAGGTGCTGCATGGGTGATTTGGGAGGACGACGAGCGCGTGT 732  
Qy 242 ValAspGlnHisProArgAlaValIleuValProSerSerLeuArgIleGlyIle 261  
Db 733 ATCAGCCCCAATACCCCGAGCGCGTGTGTGGTCCCGAGTACCTCCAGATCGGCATCAT 792

Qy	262	ProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThr	281
Db	793	CCCCGAGGGTCCACAGATTGTGTGTACTCAACAGTGGGCACAAACGACGACGACGACA	852
Qy	282	SerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerValHisHis	301
Db	853	TCGGCTTTGCATCATATTATTGGGACCTCACTGGCAATAGACGTCTCTCTGTGCACCTAC	912
Qy	302	AsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIle	321
Db	913	CATAACACGCTGCTCGGTACTCGGTTTCTCTGCTGGGCTACGGTTTCTACGGGCACTTA	972
Qy	322	IleLysAspSerGluLysLeuArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeu	341
Db	973	ATCAAGGACAGTGAAGAAACCGGTGATGGCCCTCGCCGGTATGATTTCTCAGGGTTG	1032
Qy	342	LysThrPheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis	361
Db	1033	AAGACCTTCTCTCATCAGTACTATGAGGGACACTGTCTTCTCCAGCACAGCAC	1092
Qy	362	ThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGln	381
Db	1093	ACGGTGGGATCTCCACGGGACAAATAAACCTCGCCGGGCTGGGTCTCTGTGCGAGCAG	1152
Qy	382	SerLysGlnGlnLeuGluGlnGlnLysLysAlaLeuTyrGlyLeuGluAlaGlu	401
Db	1153	AGCAAGCAACAGCTGGAGAGAGAGAGAAAGCCCTGTATGGCTCGGAACGCCGAG	1212
Qy	402	AspValGluGluTTPGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMet	421
Db	1213	GAATGGAGAGTGGCAAGTGACATGTGGGAAGTTCCTGGCCATCATGCCACCAACATG	1272
Qy	422	SerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGly	441
Db	1273	TCCTGTGCTTGTCTCGGAGCCCTGGGGCCCTGTCCCAATTGGCCACTCTGGAGATGG	1332
Qy	442	SerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIle	461
Db	1333	TCCTTCGACCTCATCTTATCCGGAAGTGTCCAGGTTCAACTTCTCGAGATTCTCTATC	1392
Qy	462	ArgHisThrAnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLys	481
Db	1393	CGGCACACGACACGAGGACGACGATTCGATTCACCTTCCTTGTGAAGTTATTCGAGTCAG	1452
Qy	482	LysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyGly	501
Db	1453	AAATTTCCATCTCACGTGCAAGCACGTGGAAGACGAGGACAAATGACTCGAAGGAACAAGAG	1512
Qy	502	LysLysArgPheGlyHisIleCysSerSerHisProSerCysCysCysThrValSerAsn	521
Db	1513	AAGCAGAAGTTTGGGAAGATCTGCAAGGACACAGACCCCTCTTGCACTTGCTCAGCCCTCAGA	1572
Qy	522	SerSerTrpAnCysAspGlyValLeuHisSerProAlaIleGluValArgValHis	541
Db	1573	AGCTCTGGAAGTTCGATGGGGAAGTCATGACACGCCCGGCATTTGAGGTGAGGTTCAC	1632
Qy	542	CysGlnLeuValArgPheAlaArgGlyIleGluGlu-----AsnPro	556
Db	1633	TGCCAGCTGGTGGCGCTCTTGTCTCGGGGAATCGAGGAAGAGTCATAGCAAGAACCCCA	1692
Qy	557	LysProAspSer	560
Db	1693	AAGCCCGAGGAGC	1704
RESULT 2			
AK052269			
LOCUS	AK052269	4248 bp	mRNA linear HTC 03-APR-2004
DEFINITION	Mus musculus 13 days embryo heart cDNA, RIKEN full-length enriched library, clone:D330016D08 product:DA59H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS) (FRAGMENT) homolog [Homo sapiens], full insert sequence.		
ACCESSION	AK052269		
VERSION	AK052269.1	GI:26342491	

# KEYWORDS

SOURCE  
Mus musculus

ORGANISM  
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mus.

# REFERENCE

AUTHORS  
Carninci, P. and Hayashizaki, Y.

TITLE  
High-efficiency full-length cDNA cloning

JOURNAL  
Meth. Enzymol. 303, 19-44 (1999)

MEDLINE  
99279253

PUBMED  
10349636

# REFERENCE

AUTHORS  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL  
Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE  
20499374

PUBMED  
11042159

# REFERENCE

AUTHORS  
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL  
Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE  
20530913

PUBMED  
11076861

# REFERENCE

AUTHORS  
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE  
Functional annotation of a full-length mouse cDNA collection

JOURNAL  
Nature 409, 685-690 (2001)

MEDLINE  
11076861

# REFERENCE

AUTHORS  
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL  
Nature 420, 563-573 (2002)

MEDLINE  
11076861

PUBMED  
11076861

# REFERENCE

AUTHORS  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, K., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akaira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE  
Direct Submission

JOURNAL  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/ Location/Qualifiers





**JOURNAL  
COMMENT**

Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-r@mail.nih.gov  
 Tissue Procurement: Lou Staudt  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLNL at:

## FEATURES

source

## ORIGIN

Alignment Scores:

Pred. No.:	1.92e-116	Length:	1059
Score:	1367.00	Matches:	276
Percent Similarity:	92.7%	Conservative:	5
Best Local Similarity:	91.0%	Mismatches:	104
Query Match:	45.19%	Indels:	8
DB:	5	Gaps:	4

US-10-631-958-11 (1-562) x B0054406 (1-1059)

Qy	158	LeuLeuValPheIleAsnProPheClyGlyysgdygldnglylyysAsgflletytrGluAarg	177
Db	15	TTACTGGTATTTATCAACCGGTTTGGAGCAAAAGCAAGCAAGCGGATATATGAAGA	74
Qy	178	LysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHis	197
Db	75	AAAGTGGCAACCACTGTTTACCCTTAGCCTCCATCACCACTGACATCATCTGTTACTTGAACAT	134
Qy	198	AlaAsnGlnAlaLysGluThrLeuTyrrGluIleAsnIleAspIlystYrAspGlyIleVal	217
Db	135	GCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGACGCGCATCGTC	194
Qy	218	CysValGlyClyAspGlyMetPheSerGluValIleuHisgIyLeuIleGlyArgThrGln	237
Db	195	TGTGTGGCGGAGATGGTATGTTACGCGAGTGTCTGCACGGTCTGATTTGGAGAGCGCAG	254
Qy	238	ArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArg	257
Db	255	AGGAGCGCGGGGTGCACGAGNACCACCCCGGGCTGTGTGTCTCCCATGTAGCCTCCGG	314
Qy	258	IleGlyIleIleProAlaGlySerThrAspCysValCysTyrrSerThrValGlyThrSer	277
Db	315	ATTGGAATCATTTCCGCGAGGTCAACGAGCTGCGGTGTGTACTCCACCGTGGCGACGAGC	374
Qy	278	AspAlaGluThr-SerAlaLeuHisIleValValGlyAspSerLeuIleAlaMetAspValSer	297
Db	375	GAGCGAAGAACTCTGGCGCTGCATATCTGTTGGGACCTCGCTGGCCATGGATGTGTC	434
Qy	298	SerValHisIleAsnSerThrLeuLeuArgTyrrSerValSerLeuLeuGlyTyrrGlyPhe	317
Db	435	TCAGTSCACCAACAGACACTCTTCGTACTTCGGTGTCCCTGTGGCTACGCGTTC	494

1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

## ORIGIN

Alignment Scores:  
 Pred. No.: 3,82e-112 Length: 1078  
 Score: 1321.00 Matches: 260  
 Percent Similarity: 92.31% Conservative: 4  
 Best Local Similarity: 90.91% Mismatches: 15  
 Query Match: 43.67% Indels: 7  
 DB: 5 Gaps: 4

US-10-631-958-11 (1-562) x BQ063738 (1-1078)

QY 115 LysArgAlaArgHisArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGlu 134  
 DB 1 AGAGAGCAGCAGCGCCGCTGGAAAGTGGCGCAGGTGACTTTCTGGTGTCAGAGGAG 60  
 QY 135 GlnLeuCysHisLeuTrpLeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArg 154  
 DB 61 CAGCTGTGTCACTTGTGGCTGCAGACCCCTCGCGAGATGCTGGAGAAGCTGACGTCAGA 120  
 QY 155 ProLysHisLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 174  
 DB 121 CCAAAGCAATTTACTGGTATTTATCAACCGTTTGGAGGAAAGGACAAAGCGGAGATA 180  
 QY 175 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleVal 194  
 DB 181 TATGAAGAAAGTGGCCACCTGTTCACTTAGCTTCATCACCCTGACATCATCGTT 240  
 QY 195 ThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAsp 214  
 DB 241 ACTGAACATGCTAATCAGCCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGAC 300  
 QY 215 GlyIleValCysValGlyLysGlyMetPheSerGluValLeuHisGlyLeuIleGly 234  
 DB 301 GGCATCGTCTGTGTGGCGGAGATGGTATGTTTCAGCGAGGTCTGCACGGTCTGATTGGG 360  
 QY 235 ArgThrGlnArgSerAlaGlyValAlaAspGlnAenHisProArgAlaValLeuValProSer 254  
 DB 361 AGCAGCAGAGAGCGCCGGGTGCACAGAACACCCCGGGCTGTGTGTCTCCCACT 420  
 QY 255 SerLeuArgIleGlyIleProAlaGlySerThrAspCysValCysTyrSerThrVal 274  
 DB 421 AGCCTCCGATTGGAATCATTCGCGAGGTCAACGGACTGGTGTGTACTTCCACCGTG 480  
 QY 275 GlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMet 294  
 DB 481 GGCACAGCGAGCGAGAAACCTCGCGCTGCATATCTGTTTGGGGACTCGCTGGCCATG 540  
 QY 295 AspValSerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGly 314  
 DB 541 GATGTCTCTAGTCCACCAACAGCACATCTCTTCGCTACTCTCGCTGCTCTGCTGGGC 600  
 QY 315 TyrGlyPheTyrGlyAspIleLysAspSerGluLysLysArgTrpLeuGlyLeuAla 334  
 DB 601 TAGCGCTTCTACGGGACATCATCAAGGACAGTGAAGAAACGGTGTGGTCTTGGC 660  
 QY 335 ArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisCysTyrGluGlyThrVal 354  
 DB 661 AGATACGACTTTTCAAGTTTAAGACCTTCTCTCCACCACTGCTATGAAGGACAGGG 720  
 QY 355 SerPheLeuProAlaGlnHisThrValGlySer---ProArgAsp-ArgLysProCysArg 373  
 DB 721 TCTTCTCTCTCCCTGCAACACACAGCGGTGGGATCTCCAGGGAATAAGGAAGCCCTGCC 780  
 QY 373 gAla---GlyCys-----PheValCysArgGlnSerLysGlnGlnLeuGluGluGlu--- 389  
 DB 781 GGGCCAGGGATGCTTTGGTTTGGTCCAGGCCAAGGCAAGCCAGCTTGGNAGGCAGGG 840

QY 390 ---GlnLysLysGala 393  
 DB 841 AGCCCAAGAAAGGCC 856

## RESULT 6

CK000755  
 LOCUS CK000755  
 DEFINITION AGENCOURT\_16363467 NIH\_MGC\_221 Homo sapiens cDNA clone  
 IMAGE:30707875 5', mRNA sequence.  
 ACCESSION CK000755  
 VERSION CK000755  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
 1 (bases 1 to 797)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: cgapbs-remail.nih.gov

Tissue Procurement: James Martin, University of Iowa  
 cDNA Library Preparation: M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ILLNL at:  
 http://image.llnl.gov  
 Plate: NDAM1073 row: c column: 20  
 High quality sequence stop: 656.

## FEATURES

Location/Qualifiers  
 1..797  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30707875"  
 /lab\_host="DH10B Tona"  
 /clone\_lib="NIH\_MGC\_221"  
 /note="Organ: mixed; Vector: pYX-Asc; Site: 1: EcoRI;  
 Site 2: NotI; Library is oligo-dT primed and directionally  
 cloned. Denatured RNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with oligo-dT  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to mRNA size fraction, ligated with  
 EcoR I adaptor, digested with Not I and then cloned  
 directionally into pYX-Asc vector. Average insert size  
 4-5Kb. Adaptors 5' (AATCGGACGAGG)3' and 5'  
 (CTCTGTCGG)3'. 3' linker sequence - GCGGCCGTCGAGGCC T18.  
 Sequencing primers 3' end: T3 promoter primer 5'  
 (ATTAAACCTCACTAAAGGA)3', 5' End: T7 promoter primer 5'  
 (TAATAGACTCACTATAGG)3'. Library was constructed in the  
 laboratory of M. Bento Soares. Note: this is a NIH\_MGC  
 Library"

## ORIGIN

Alignment Scores:  
 Pred. No.: 2,42e-110 Length: 797  
 Score: 1299.50 Matches: 250  
 Percent Similarity: 96.54% Conservative: 1  
 Best Local Similarity: 96.15% Mismatches: 8  
 Query Match: 42.96% Indels: 1  
 DB: 7 Gaps: 1

US-10-631-958-11 (1-562) x CK000755 (1-797)

QY 177 ArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleValThrGlu 196  
 DB 9 CGGCACGAGGACCACTGTTCCCTTCCATCACCCTGACATCATCTGTTACTGAA 68



```
Qy 197 HisAlaAsnGlnAlaLysGluThrLeuTyrGluLeuAsnIleAspLysTyrAspGlyLe 216
Db 69 CATGCTAATCAGGCCAAGAGACTCTGTATGAGATTAAATAGACAAATACGCGGCATC 128
Qy 217 ValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThr 236
Db 129 GTCTGTGTGCGCGGAGATGATGTTTCAGCGAGGTGTCACGGTCTGTATGGGAGACG 188
Qy 237 GlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeu 256
Db 189 CAGAGAGCGCGGGGTGCACAGAACACCCCGGGCTGTGCTGCTCCCGATGAGCCTC 248
Qy 257 ArgIleGlyIleIleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThr 276
Db 249 CGGATTGGAATCATTCCTCCGCGAGGTCAACGCACTGCGGTGCTTACTCCACCGTGGGCACC 308
Qy 277 SerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspVal 296
Db 309 AGCGACGCAGAAACCTCGCGCTGCATATCGTTGTTGGGAGACTCGCTGGCCATGGATGTG 368
Qy 297 SerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGly 316
Db 369 TCCTCAGTCCACCAACAGACACTCTCTCGCTACTCCGTCCTCCGTCCTGGCTACGGC 428
Qy 317 PheTyrGlyAspIleIleLysAspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyr 336
Db 429 TTCTACGGGACATCATCAAGGACAGTGAAGAAGACGGTGTGGTCTTGGCCAGATAC 488
Qy 337 AspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPhe 356
Db 489 GACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTCTTC 548
Qy 357 LeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCys 376
Db 549 CTCTCTGCACAAACACACGGTGAGATCTCCAAGGATAGGAAGCCCTCGCGGGCAGGATGC 608
Qy 377 PheValCysArgGlnSerLysGlnIleLeuGluGluGlnLysLysAlaLeuTyrGly 396
Db 609 TTTGTTTCAGCGAAAGCAAGCAGCAGCTGGAGGAGGAGCAGAAAGCACTGTATGGT 668
Qy 397 LeuGluAlaAlaGluAspValGluGluTyrGlnValValCysGlyLysPheLeuAlaIle 416
Db 669 TTGGAAGCTGCGGAGGACGTGGAGAGTGGNCAGTCTGTCTGGGAAATTTCTCGGGCNA 728
Qy 417 AsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAla 436
Db 729 TCATGCACAAACATGCTCTGTGCTGTGTCGCCGAGCCCA---GGGGCTCTCCCGGTGCC 785

RESULT 7
BOQ57191 1047 bp mRNA linear EST 29-MAR-2002
LOCUS AGENCOURT_6769595 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5812382
DEFINITION 5', mRNA sequence.
ACCESSION BOQ57191
VERSION BOQ57191.1 GI:19816531
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1047)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
```

```
Plate: LLC2062 row: 1 column: 15
High quality sequence stop: 535.
FEATURES
source
1..1047
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5812382"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
```

ALIGNMENT SCORES

Pred. No.:	1.03e-108	Length:	1047
Score:	1284.00	Matches:	274
Percent Similarity:	91.67%	Conservative:	12
Best Local Similarity:	87.82%	Mismatches:	13
Query Match:	42.45%	Indels:	13
DB:	5	Gaps:	2

US-10-631-958-11 (1-562) x BQ057191 (1-1047)

Qy	158	LeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArg	177
Db	13	TTACTGCTATTATTAACACCGTTTGGAGGAAGAAGCAAGGCAAGCGATATATGAAGA	72
Qy	178	LysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleValThrGluHis	197
Db	73	AAAGTGGCACACCTGTTTCACTTAGCTCCATCACCCTGACATCATCTGTTACTGAACAT	132
Qy	198	AlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleVal	217
Db	133	GCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGCGCATCGTC	192
Qy	218	CysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGln	237
Db	193	TGTGTGCGGAGATGATGTTGTTGAGGTGCTGCACGGTCTGATTTGGAGACCGCAG	252
Qy	238	ArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArg	257
Db	253	AGGAGCGCGGGGTGCGACCAAGAACCCCGGGCTGTGCTGCTCCCGTAGCTCCGG	312
Qy	258	IleGlyIleIleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSer	277
Db	313	ATTGGAATCATTTCCGCGAGGTCAACGCACTGCGGTGTGTACTCCACCGTGGGACCGAC	372
Qy	278	AspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSer	297
Db	373	GACCAAGAAACCTCGCGCTGCATATGTTGTTGGGACTCGCTGGCCATGGATGTGTCC	432
Qy	298	SerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPhe	317
Db	433	TCAGTCCACCAACACAGCACACTCTCTCGCTACTCTCGTCTGCTGCTGGGTACGGCTTC	492
Qy	318	TyrGlyAspIleIleLysAspSerGluLysLysArgTyrPheGlyLeuAlaArgTyrAsp	337
Db	493	TACGGGACATCATCAAGGACAGTGAAGAAGACGGTGGTGGGTCTTGGCAGATACGAC	552
Qy	338	PheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGlu-GlyThrValSerPheLe	357
Db	553	TTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAGGGGACAGATGCTCTCCT	612
Qy	357	uProAlaGlnHisThrValGlySer-ProArgAspArgLysProCys-ArgAlaGlyCys	376



Db 613 CCTCGACACACACGGTGGATCTCCCAAGGATAGGAAACCTGCGCGGCACGATGC 672  
 QY 377 PheVal-CysArgGlnSerLysGlnLeuGluGlu--GlnLysLysAlaLeuTyr 395  
 Db 673 TTTGTTGTCGCGCAAGCAAGCAGCAGCTGGAGGAGGAGCCCAAGAAACCACTTGTAT 732  
 QY 396 GlyLeuGluAlaAlaGluAspValGluGluTTPGlnValValCys-GlyLysPheLeuAl 415  
 Db 733 GTTTTGAACCTGCGGGAACCGCGGAGGAGTGGCAATCTTCTGGGGGAGTGTCTGGC 792  
 QY 415 alleAsnAlaThrAsnMetSerCys-AlaCysArgArgSerProArgGlyLeuSer--P 434  
 Db 793 CATCATGCCCAACCTGCTCTGGTGTCTTGTGCGGGAACCCCGAGGGGCTTCTCCCC 852  
 QY 434 roAlaAlaHisLeuGlyAsp---GlySerSerAspLeuLe-LeuLe-ArgLysCys-S 452  
 Db 853 CGGCTGCCCAATTTGGGGAAGAGGCTCTTCTGAACTCTCTCTCTCTCTCTCTCTCTCT 912  
 QY 452 erArgPheAsnPhe 456  
 Db 913 CCCGTTTCATTTT 926

CK000525 713 bp mRNA linear EST 26-NOV-2003  
 AGENCOURT\_16369000 NIH\_MGC\_221 Homo sapiens cDNA clone  
 IMAGE:30708597 5', mRNA sequence.  
 CK000525  
 CK000525.1 GI:38526559  
 EST.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 713)  
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>  
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Unpublished (1999)  
 Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: c9apbs-r@mail.nih.gov  
 Tissue Procurement: James Martin, University of Iowa  
 cDNA Library Preparation: M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: NDAM1075 row: a column: 22  
 High quality sequence stop: 689.  
 Location/Qualifiers

FEATURES  
 source  
 1..713  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30708597"  
 /lab\_host="DH10B Tona"  
 /clone\_lib="NIH\_MGC\_221"  
 /note="Organ: mixed; Vector: pYX-Aec; Site 1: EcoRI;  
 Site 2: NotI; Library is oligo-dT primed and directionally  
 cloned. Denatured RNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with oligo-dT  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to mRNA size fraction, ligated with  
 EcoR I adaptor, digested with Not I and then cloned  
 directionally into pYX-Aec vector. Average insert size  
 4-5Kb. Adaptors 5'(AATTCGCACGAGG)3' and 5'd  
 (CTCGTGGCG)3'. 3' Linker sequence - GCGGCGCTGAGAGCC T18.  
 Sequencing primers 3'end: T3 promoter primer 5'd  
 (ATTAAACCTCACTAAGGA)3'. 5' End: T7 promoter primer 5'd

(TATACGACTCACTATAGG)3'. Library was constructed in the  
 library of M. Bento Soares. Note: this is a NIH\_MGC  
 library"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 3.92e-108 Length: 713  
 Score: 1275.00 Matches: 235  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 42.15% Indels: 0  
 DB: 7 Gaps: 0

US-10-631-958-11 (1-562) x CK000525 (1-713)

QY 302 AsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIle 321  
 Db 9 AACAGCACACTCTTCGCTACTCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGGACATC 68  
 QY 322 IleLysAspSerGluLysLysArgTyrLeuAlaArgTyrAspPheSerGlyLeu 341  
 Db 69 ATCAAGGACAGTGAAGAAACGGTGTGGTCTTGCAGATACGACTTTTCAGTTTA 128  
 QY 342 LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis 361  
 Db 129 AAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCCTCCCTGCACAACAC 188  
 QY 362 ThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGln 381  
 Db 189 ACGTGGGATCTCCAAAGGATAGGAGCCCTGCCGGGAGGATGCTTTGTTTCAGGCCAA 248  
 QY 382 SerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaGlu 401  
 Db 249 AGCAAGCAGCAGCTGGAGGAGGAGCAGAAAGACACTGTATGTTTGGAAAGCTCGGAG 308  
 QY 402 AspValGluGluTTPGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMet 421  
 Db 309 GACGTGGAGGAGTGGCAAGTCTGTGGGAAGTTTCTGGCCATCATGCCACAACATG 368  
 QY 422 SerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGly 441  
 Db 369 TCCTGTGTTCTGCCGAGGAGCCCGAGGGCTCTCCCGGGCTGCCACTTGGGAGACGG 428  
 QY 442 SerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIle 461  
 Db 429 TCTTCTGACCTCATCTCATCCGGAATGCTCCAGTTTCAATTTTCTGAGATTCTCATC 488  
 QY 462 ArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLys 481  
 Db 489 AGGCACACCAACACGAGGACAGTTCAGTTTCACTTTTGTGAAGTTTATCGCGTCAAG 548  
 QY 482 LysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyGly 501  
 Db 549 AATTTCCAGTTTACGTCGAAGCACATGAGGATGAGGACACGACCTCAAGAGGGGGGG 608  
 QY 502 LysLysArgPheGlyHisIleCysSerSerHisProSerCysCysCysThrValSerAsn 521  
 Db 609 AAGAAGCGCTTTGGGCACATTTGCAGCAGCAGCCCTCTCTGCTGTCGACCGTCTCCAAC 668  
 QY 522 SerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIle 536  
 Db 669 AGCTCTTGAAGTCTGAGGGGAGGTCTCTGCACAGTCTCTGCATC 713

RESULT 9

LOCUS

CD655311 820 bp mRNA linear EST 18-JUN-2003  
 AGENCOURT\_14552675 NIA Human HI Embryonic Stem Cell cDNA Library  
 (Long) Homo sapiens cDNA clone IMAGE:30426593 5', mRNA sequence.

ACCESSION

CD655311

VERSION

CD655311.1

GI:31895467

EST.

KEYWORDS

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 820)  
NIH-MGC <http://mgc.nci.nih.gov/>  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NCI  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Irene Ginis and Mahendra Rao, NIA  
CDNA Library Preparation: Yulan Piao and Minoru Ko  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC c lone distribution information  
can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: NDAM512 row: k column: 18  
High quality sequence stop: 673.

FEATURES  
source

1. 820  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30426593"  
/tissue\_type="Embryonic stem cells"  
/cell\_line="WA01"  
/lab\_host="DH10B (T1 phase-resistant)"  
/clone\_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"  
/note="Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI;  
This is a long-transcript enriched cDNA library (Genome Res 11: 1553-1558 (2001). [PMID: 11544199]) from WA01 cell line. Undifferentiated human ES cell line WA01/H1 was obtained from WiCell Research Institute, Inc., Madison, WI, cultured according to their instructions, on MEK feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UNF, TERT, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1, TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days after plating), the ES cells from 4 X 6cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with Trizol reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID:11544199]) Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTCTAGATCGGAGCGCGCCCTTTT-3'] from 3.4g of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lona-linker LL-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by Centricron-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S for 25 cycles. The products were purified by phenol/chloroform extraction and Centricron-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-Sport6 plasmid vector. The average insert size is about 3.6kb."

ORIGIN

Alignment Scores:  
Pred. No.: 1,228-106 Length: 820  
Score: 1260.00 Matches: 250  
Percent Similarity: 93.12% Conservative: 7  
Best Local Similarity: 90.58% Mismatches: 12  
Query Match: 41.65% Indels: 8  
Dbs: 6 Gaps: 5

US-10-631-958-11 (1-562) x CD655311 (1-820)

Qy	272	SerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSer	291
Db	12	TCACCCGTGGGACCGAGGACGAGAAACCTCGGCGCTGCATATCGTTGTTGGGACTCG	71
Qy	292	LeuAlaMetAspValSerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSer	311
Db	72	CTGGCCATGGATGTCTCAGTCCACCAACAGCACACTCCTTCGCTACTCCGCTGTCC	131
Qy	312	LeuLeuGlyTyrGlyPheTyrGlyAspIleIleLeuSerSerGluLeuValAspTyrLeu	331
Db	132	CTGCTGGGCTACGGCTCTTACGGGGAATCATCAAGGACAGTGGAGAAACGGTGGTTG	191
Qy	332	GlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGlu	351
Db	192	GGTCTTCCAGATACGACTTTTTCAGGTTAGAGACCTTCTCTCCCAACCACTGCTATGAA	251
Qy	352	GlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysPro	371
Db	252	GGGACAGTGTCTTCTCTCCCTGCACAAACACGCGTGGGATCTCCAGGAGTAGGAAGCC	311
Qy	372	CysArgAlaGlyCysPheValCysArgGlnSerLysGlnLeuGluGluGlnLys	391
Db	312	TGCGGGGAGGATGCTTTTTCAGGCAAGCAAGCAGCAGCTGGAGGAGGAGCAGCAAG	371
Qy	392	LysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGluTyrGlnValValCysGly	411
Db	372	AAACCACTGTATGTTGGTGGAGCTCGGAGGAGCTGGAGGAGTGCCTGCTGTGGG	431
Qy	412	LysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGly	431
Db	432	AAAGTTCTGGCCATCAATATGCCAACATGTCCTGTGCTGTGTCGCCGAGGCCAGGGGC	491
Qy	432	LeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIleLeuLeuValCys	451
Db	492	CTCTCCCGGCTGCCACTTGGGAGAGCGGTCTTCTGACCTCATCTCATCCGGAATGC	551
Qy	452	SerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsnGlnGlnAspGlnPheAsp	471
Db	552	TCCAGGTTCAATTTCTGAGATTTCTCATCAGGCACACCAACAGCAGGACCACTTTGAC	611
Qy	472	PheThrPheValGluValTyrArgValLysLysPheGlnPheThrSerLysHisMetGlu	491
Db	612	TTCACTTTGTTGAAGTTTATCGGTCGAAGAAATTCAGATTTCAGTCCGAAGCACATGGAG	671
Qy	492	AspGluAspSerAspLeuLysGluGlyLysLysArgPheGlyHisIleCysSerSer	511
Db	672	GATGAGGACGCGACCTCAAGGAGGNGGAGAGCGCTGGGCACAT---TGCAGGACG	728
Qy	512	HisProSerCysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeu	531
Db	729	---ACCTCTGCTCTGCACG---TCTCACAGCTCTG---GACTCGCAGCGGAGTCTG---	777
Qy	532	HisSerProAla---IleGluValArgValHisCysGlnLeuValArg	546
Db	778	-----CGAGCTGATCATGTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG	819
RESULT 10			
CF135528			
LOCUS	CF135528	732 bp	mRNA linear EST 09-SEP-2003
DEFINITION	UI-HF-BNO-amf-g-10-0-UI.r1 NIH MGC_50 Homo sapiens cDNA clone		
	IMAGE:3090211 5', mRNA sequence.		
ACCESSION	CF135528		
VERSION	CF135528.1	GI:33250972	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 732)		
AUTHORS	Bonaldo, M.F., Lennon, G. and Soares, M.B.		
TITLE	Normalization and subtraction: two approaches to facilitate gene		
	discovery		

JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Genome Res. 6 (9), 791-806 (1996)  
97044477  
8889548  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: Louis Staudt  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
http://genome.uiowa.edu/distribution/humanfl.html  
Seq primer: pYX-5.

FEATURES  
source

1. .732  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3090211"  
/tissue\_type="lymph"  
/cell\_type="germinal center B cells"  
/cell\_line="MG85"  
/lab\_host="DH10B (LTI)"  
/clone\_lib="NIH MGC 50"  
/note="Vector: pYT73-Pac; Site 1: NotI; Site 2: Eco RI;  
Constructed from size fractionated cytoplasmic mRNA  
(3.5-4.4kb). Directionally cloned. Cells provided by  
Louis M. Staudt, Ph.D. Library preparation by Maria de  
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

ORIGIN

Alignment Scores:  
Pred. No.: 3,16e-105 Length: 732  
Score: 1244.00 Matches: 231  
Percent Similarity: 99.57% Conservative: 0  
Best Local Similarity: 99.57% Mismatches: 0  
Query Match: 41.12% Indels: 1  
DB: 7 Gaps: 0

US-10-631-958-11 (1-562) x CF135528 (1-732)

QY 332 GlyLeuAlaArgTyxAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrglu 351  
Db 9 GGTCTTGCCAGATACGACTTTTCAGGTTTAAGACCTTCTCTCCACCACTGCTATGAA 68  
QY 352 GlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysPro 371  
Db 69 GGGACAGTGTCTCTCTCCCTCACACACACGGTGGGATCTCCAGGGATAGGAAGCCC 128  
QY 372 CysArgAlaGlyCysePheValCyseArgGlnSerLysGlnGlnLeuGluGluGlnLys 391  
Db 129 TCCCGGGCAGGAGTCTTTGTCAGGCAAGCAAGCAGCAGCTGGAGGAGGAGGAGGAG 188  
QY 392 LysAlaLeuTyxGlyLeuGluAlaAlaGluAspValGluGluTrpGlnValValCysGly 411  
Db 189 AAAGCACTGTATGGTTTGAAGCTCGGAGGAGCTGGAGGAGTGGCAGTCTCTGTGGG 248  
QY 412 LysPheLeuAlaIleAsnAlaThrAsnMetSerCyseAlaCyseArgSerProArgGly 431  
Db 249 AAGTTTCTGGCCATCAATGCCAACATGTCTGTGCTGTGTGCGCGGAGCCCGAGGCGC 308  
QY 432 LeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIleLeuArgLysCys 451  
Db 309 CTTCTCCCGGCTGCCACTTGGAGACGGGTCTTCGACCTCATCTCATCCGAAATGC 368  
QY 452 SerArgPheAsnPhLeuLeuArgPheLeuIleArgHisThrAsnGlnGlnAspGlnPheAsp 471  
Db 369 TCCAGGTTCAATTTCTGAGATTTCTCATCAGGCACACCAACAGCAGGACCACTTTCAC 428

QY 472 PheThrPheValGluValTyxArgValLysLysPheGlnPheThrSerLysHisMetGlu 491  
Db 429 TTCACCTTTTGTGAAGTTTATCGCGCTCAAGAAATTCAGATTTACGTCAAGACATGGAG 488  
QY 492 AspGluAspSerAspLeuLys-GluGlyGlyLysLysArgPheGlyHisIleCysSerSe 511  
Db 489 GATGAGGACAGGACCTCAGGGAGGGGGGAGGAGGCTTTGGGCACATTTGGACGAG 548  
QY 511 rHisProSerCysCyseThrValSerAsnSerSerTrpAsnCysAspGlyGluVal 531  
Db 549 CCACCCCT 608  
QY 531 uHisSerProAlaIleGluValAlaGlnValHisCysGlnLeuValAlaGlnPheAlaArgG 551  
Db 609 GCACAGCCCTGTCATCGAGGTGAGAGTCCAGTCCAGCCAGCTGGTTCGACTCTTTTGCACGAGG 668  
QY 551 yileGluGluAsnProLysProAspSerHisSer 562  
Db 669 GATTGAAGAGAAATCCGAAGCCAGACTCACACAGC 702

RESULT 11

LOCUS CB246749  
DEFINITION UI-M-F10-cdx-b-10-0-UI.r1 NIH\_BMAP\_F10 Mus musculus cDNA clone  
IMAGE:5835595 5', mRNA sequence.  
ACCESSION CB246749  
VERSION 1 GI:28368393  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 758)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project  
(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1. .758  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6835595"  
/tissue\_type="whole brain"  
/dev\_stages="embryo 12.5dpc"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_lib="NIH\_BMAP\_F10"

/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;  
Site 2: Not I; The library was constructed according  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT  
primer containing a Not I site. Double strand cDNA was  
size selected according to mRNA size fraction, ligated  
with EcoR I adaptor, digested with NotI and then cloned  
directionally into pYX-Asc vector. The library tag  
sequence located between the Not I site and the polyA tail  
is CAGCCACGAC. This library was created for the University  
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the  
Developing Mouse Nervous System', supported by National  
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,

SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE		1 (bases 1 to 653)	
AUTHORS		Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.	
TITLE		EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., et al.)	
JOURNAL		Unpublished (2003)	
COMMENT		Contact: MIPS	
MIPS			
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany			
This is the 5' sequence of the clone insert			
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.			
No sl sequence available.			
This clone (DKFZp781l11183) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.			
FEATURES		Location/Qualifiers	
source		1..653	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="DKFZp781l11183"	
		/dev_stage="adult"	
		/lab_host="DH10B"	
		/clone_lib="781 (synonym: hlcc4)"	
		/note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB; cDNA-collection"	
ORIGIN			
Alignment Scores:			
Pred. No.: 1.98e-99 Length: 653			
Score: 1181.00 Matches: 217			
Percent Similarity: 100.00% Conservative: 0			
Best Local Similarity: 100.00% Mismatches: 0			
Query Match: 39.04% Indels: 0			
DB: 5 Gaps: 0			
US-10-631-958-11 (1-562) x BX952302 (1-653)			
Qy	322	IleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeu	341
Db	2	ATCAAGGACAGTGAGAAGAAACGGTGGTGGTCTGCCAGATACGACTTTCAGGTTTA	61
Qy	342	LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis	361
Db	62	AAGACCTTCTCTCCCACTGCTATGAAGGAGCAGTGTCTCTCTCCCTGCACAAAC	121
Qy	362	ThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGln	381
Db	122	ACGTGGGATCTCCAAGGATAGGAAGCCCTGCCGGCAGGATGCTTGTGTTCAGGCAA	181
Qy	382	SerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGlu	401
Db	182	AGCAAGCAGCAGCTGGAGGAGGAGCAGAGAAGCACTGTATGTGTTGGAGCTCGGAG	241
Qy	402	AspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMet	421
Db	242	GACGTGGAGGATGGCAAGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATG	301
Qy	422	SerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGly	441
Db	302	TCCTGTGCTGTGTCGGGAGGCCCGAGGGCCCTCTCCCGGCTGCCACTTGGGAGACGGG	361
Qy	442	SerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIle	461
Db	362	TCTTCTGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTCTCATC	421
Qy	462	ArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgVallys	481



/clone.lib="Lupski\_dorsal\_root\_ganglion"  
/note="vector: pCMV-SPORT6 (Life Technologies); Site 1:  
NotI; Site 2: SalI; cDNA made by oligo-dT priming.  
Directionally cloned using the following adaptors:  
5'-TCGACCCACGGCTCGG-3' and  
5'-GACTAGTCTAGATCGGAGCGGCCCT(15)-3'. Size selected >  
1 kb for average insert length 1.7 kb. This is a primary  
library, non-amplified. Library constructed by Life  
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
College of Medicine) and is available through Life  
Technologies."

## ORIGIN

Alignment Scores:  
Pred. No.: 3.1e-92 Length: 1017  
Score: 1107.00 Matches: 223  
Percent Similarity: 86.54% Conservatives: 2  
Best Local Similarity: 85.77% Mismatches: 6  
Query Match: 36.60% Indels: 30  
DB: 5 Gaps: 3

US-10-631-958-11 (1-562) x B0879739 (1-1017)

Qy 7 ProAlaProLeuGlyValArgAlaProProAlaTrpArgThrSerProAlaAlaGluMet 26  
Db 2 CCGGCGCCCTCGGCGTCCGCGCCCGCCAGCTGGCGAGCAGCCGCGCGGAGATG 61

Qy 27 GlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValValGlnArgCys 46  
Db 62 GGGGCGAGCGGGCGGGCGGAGCCGCTGCATCCGTGCTGGGTGAACGACGACGCGTC 121

Qy 47 AlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyProGly 66  
Db 122 GCGGTAGCTTGGAGCCCGCGGGCTCTGTGCGTGTGGCGAGCCCGGGCCCGGA 181

Qy 67 AlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluLeuLeuAlaValGlu 86  
Db 182 GCGGCGCCCGCGCGCGGATGCTCTGTCTGTCTGTATCTGAGATCATCGCGTTGAG 241

Qy 87 GluThrAspValHisGlyValHisGlnGlySerGlyValTrpGlnValMetGluValPro 106  
Db 242 GAACAGACGTTACGGGAAACATCAGCGCAGTGGAAATGGCAGAAATGGAAAGCCT 301

Qy 107 TyrAlaPheThrValHisCysValHisArgAlaArgArgHisArgTrpLysTrpAlaGln 126  
Db 302 TACGCTTTTACAGTTCACTGTGTAAAGAGACGACGCGCCGCTGGAAGTGGCGCAG 361

Qy 127 ValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArgGlu 146  
Db 362 GTGACTTCTGTGTCCAGAGAGCAGCTGTGTCACTTGTGCTGCAGACCTGCGGAG 421

Qy 147 MetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheLeuAsnProPheGly 166  
Db 422 ATGCTGGAGAGCTGACGCTCAGACCAAGACATTTACTGGTATTATCAAC----- 472

Qy 167 GlyLys-GlyGlnGlyLysArgGlyTyrGluArgLysValAlaProLeuPheThrLeuAl 186  
Db 473 ---RAAGGACAGCAGCAGCGGATATATGACAGAAAGTGGCACCACCTTTCACCTTAGC 529

Qy 186 aSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeuTy 206  
Db 530 CTCCATCACCACTGACATCATCGC----- 553

Qy 206 rGluLeuAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPheSe 226  
Db 554 -----ATCGTCTGTCTGCGCGGAGATGATGATTTTCAG 585

Qy 226 rGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAla-GlyValAspGlnAsnH 246  
Db 586 CGAGGTGTCGACCGGCTGATTGGAGGACCCAAAGAGAGCGCGGGGTTCGACAGAAC 645

Qy 246 isProArgAlaValLeuValProSer----SerLeuArgIle-GlyIleIlePro 262  
|||||

Db 646 CCCCCCGGGCTGTGCTGGCTCCCAATAGCCCTTCCGGATTGGGAATCATTTCCC 699

RESULT 15  
CK603033

LOCUS CK603033 812 bp mRNA linear EST 22-JAN-2004

DEFINITION AGENCOURT 17899852 NIH MGC\_234 Rattus norvegicus cDNA clone  
IMAGE:7193195 5', mRNA sequence.

ACCESSION CK603033

VERSION CK603033.1 GI:41116352

KEYWORDS EST

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 812)  
NIH-MGC <http://mgi.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgapps-remail.nih.gov  
Tissue Procurement: Howard Jacobs  
cDNA Library Preparation: Express Genomics  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
DNA distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1AM15055 row: b column: 09  
High quality sequence stop: 657.

FEATURES  
Location/Qualifiers  
1..812  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/clone="IMAGE:7193195"  
/tissue\_type="heart, pooled"  
/lab\_host="DH10B Tona"  
/clone\_lib="NIH\_MGC\_234"  
/note="Organ: heart; Vector: pExpress-1; Site 1: EcoRV;  
Site 2: NotI; RNA obtained from pooled heart tissue from a  
mix of male and female animals at 8 wk old. Tissues were  
snap-frozen and kept at -80C for two days before RNA  
extraction and purification (Tri-reagent method). cDNA was  
primed using oligo-dT primer:  
5'-pGACTAGTCTAGATCGGAGCGGCCCT(15)-3' and cloned into  
the EcoRV/NotI sites of pExpress-1. Size-selection >1.4kb  
resulted in an average insert size of 2.2 kb. This primary  
library is normalized (non-normalized primary library is  
NIH\_MGC\_233) and was constructed by Express Genomics  
(Frederick, MD). Note: this is a NIH\_MGC library."

Alignment Scores:  
Pred. No.: 6.53e-88 Length: 812  
Score: 1059.00 Matches: 209  
Percent Similarity: 85.13% Conservatives: 20  
Best Local Similarity: 77.70% Mismatches: 35  
Query Match: 35.01% Indels: 6  
DB: 7 Gaps: 4

US-10-631-958-11 (1-562) x CK603033 (1-812)

Qy 21 SerProAlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrp 40  
Db 3 GCGCGCGCGCGC---ATGGCGCAATGGCGCGCGCGCGCGCTGCAATCGTGTGTGG 59

Qy 41 ValLysGlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrp 60  
Db 60 GTGAAGCAGCAACGCTGTGCGCTGAGCTTGGAGCCCGCGCGGCTCTGCTACGCTGTGG 119

```
QY 61 ArgSerProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSer 80
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
120 CGGACCCCGAGCCCGGGCCCTCGCGGCCCGATCCCGATGCTACTCGGTACCAAGTGTC 179
QY 81 GluIleIleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrp 100
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
180 GAGATCATCGCGTTGAGGAAAGAGATCCACCGAAGAACAGTCTCCAAATGGCCGATGG 239
QY 101 GlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgHis 120
Db ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
240 CATAGATGGAAATCCGTTTGCAATTCACAGTCCACTGCTGTGAAGCGAGCTCGACACCAC 299
QY 121 ArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuLysHisLeuTrp 140
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
300 CGCTGGAAGTGGGCACGGGTGACCTTCTGGAGCGCCGACGAGCAGCTGTGCAACTGTGG 359
QY 141 LeuGlnThrLeuArgGluMetLeuLysLeuThrSerArgProLysHisLeuLeuVal 160
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
360 CTGCAGACCTCCGAGAGCTGCTGGAGAACTGACTTCAAGACCGAAGCAITTTGCTGGTA 419
QY 161 PheIleAsnProPheGlyGlyGlyGlnGlyLysArgIleTyrGluArgLysValAla 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
420 TTTATCAACCTTTCGGAGGGAAGGCCAGGGCAAGCACATCTATGAAAAAAGTGGA 479
QY 181 ProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGln 200
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
480 CCTTTGTTTCACTTGGCTTCCATCACCACCTGAGATCATCATTTACTGAGCATGCCAATCAA 539
QY 201 AlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGly 220
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
540 GCCAAGGAGACTTTATAGAGATCAACACAGACAGCTATGACGGCATCGTGTGTGTGT 599
QY 221 GlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAla 240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
600 GGGGACGCATGTTTCAGCGAGGTGCTGCACGGGTAAATGGGAAGACCGACAGAGCGCT 659
QY 241 GlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIle 260
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
660 GGTGTGACCCCAATCACCCCGAGCCGCTGCTGTGTGCCAGTACCCT--AGGATCGGCATC 717
QY 261 IleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGlu 280
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
718 ATCCGGCAGGGTCACAGA---TGTGTGTC---TCTCAGGTGGCACAACG-----AGC 765
QY 281 ThrSerAlaLeuHisIleValValGly 289
Db ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
766 AGACGTCGCTTACCATCATATGGGA 792
```

Search completed: September 6, 2005, 20:30:03  
Job time : 5431.27 secs



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 04:10:49 ; Search time 24.8463 Seconds  
(without alignments)  
2176.332 Million cell updates/sec

Title: US-10-631-958-11  
Perfect score: 3025  
Sequence: 1 HEAANGPAPLGVRAAPPART.....QLVRLFARGIENPKDSHS 562

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79: \*  
1: Pir1: \*  
2: Pir2: \*  
3: Pir3: \*  
4: Pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	567.5	18.8	549	2 T33517	hypothetical prote
2	417.5	13.8	1240	2 T05162	hypothetical prote
3	291.5	9.6	458	2 T38776	hypothetical prote
4	248.5	8.2	687	2 S51398	hypothetical prote
5	245.5	8.1	473	2 T19707	hypothetical prote
6	210.5	7.0	624	2 S67059	hypothetical prote
7	189.5	6.3	310	2 AG1665	hypothetical prote
8	176	5.8	310	2 A11293	hypothetical prote
9	147	4.9	303	2 F69795	conserved hypothet
10	136.5	4.5	306	2 AH1769	conserved hypothet
11	136	4.5	732	2 T16422	hypothetical prote
12	130.5	4.3	309	2 AH1528	conserved hypothet
13	129	4.3	295	2 A83894	hypothetical prote
14	126.5	4.2	306	2 AE1394	conserved hypothet
15	125.5	4.1	295	2 D83734	hypothetical prote
16	125.5	4.1	309	2 AF1171	conserved hypothet
17	122.5	4.0	433	2 S75948	hypothetical prote
18	115.5	3.8	294	2 G95120	conserved hypothet
19	114	3.8	364	2 F84898	hypothetical prote
20	112	3.7	345	2 E69678	involved in polyke
21	111.5	3.7	311	2 C97990	hypothetical prote
22	105	3.5	333	2 F71006	hypothetical prote
23	103.5	3.4	1028	2 A96719	hypothetical prote
24	102.5	3.4	315	2 A89978	conserved hypothet
25	102	3.4	297	2 F69595	multidrug resistan
26	102	3.4	343	2 C97183	cathepsin Q (EC 3.
27	100	3.3	650	1 JCI450	fibroblast growth
28	99	3.3	1555	2 T18688	hypothetical prote
29	99	3.3	1973	2 G89608	protein B0272.5 [i

Query Match 18.8%; Score 567.5; DB 2; Length 549;  
Best Local Similarity 30.8%; Pred. No. 2.1e-38;  
Matches 139; Conservative 91; Mismatches 160; Indels 61; Gaps 13;  
QY 114 VKARRHRKWAQ--VTFWCPPEQLCHLWQLTREMLEKLTSPKHLVFINPFGKGQG 171  
Db 123 VYKDKQKRLKQIPVIFYTTSER--DYVHSLDTTLRRVKNRPKNIIIFINPFGNGKA 180  
QY 172 KRIYERKVAFLTLA-SITTDIIIVTEHANQAKELYEINIDKY---DGIVCVGGDMFSE 227  
Db 181 QKIFKDNVDAFFWLTPLGLRYKVLTERANHARDYIVEMPPQWSAIDGLSVSGDGLFNE 240  
QY 228 VHLGLIGHTQSAGVDQNHPRAVLPSSLRIGIIPAGSTDCVCYSTVGTSTDAETSALHIV 287  
Db 241 LLSGALLRTQDAGRINDNPSHSLVTPHIRFGIIGAGSANSIVSTVHTNDHATSAVHIA 300  
QY 288 VGDSLAMDVSVHNSLTLLRYSVSLGFGVDGIIDKSEKKRWGLGLARYDFSLGKTFLSH 347  
Db 301 IGSECNVDVCTVHOHQKLIRISANAISYGMVGDLVDLRDSEYRCGLGPIRYQWSALRTTIRH 360  
QY 348 HCYEYGTVSFLPAQHTVGSPPDR-KPCRACGFCVCRSQKQLEBEQKALYLEAAEDVBEW 406  
Db 361 PIYRGWQFSLSHKENVNPDKQLPCLPCPCVMK-----PQGNDDKYDYHWAH----EFT 411  
QY 407 QVVGCKFLAIATNWSACRRSPRGLSPAHLGGSSDLILIRKCSRNFNRLFLRHT-- 464  
Db 412 HVICCVIPTVTPF-----TPYGLAPFTGIGDGLTDLALVPRISRPHNQFMKRVAMY 463  
QY 465 --NQODQDFTFVEYVRYVKFQFTSKHMEDESDLKEGKKRFGHICSSHSCCCTVNS 522  
Db 464 GKGQLYELD-PSLNCYRVTKWSY-----QPDADQEDPG----- 495

hypothetical prote  
protein C23F12.1 l  
hypothetical prote  
probable membrane  
conserved hypothet  
hypothetical prote  
conserved hypothet  
hypothetical prote  
hypothetical prote  
hypothetical prote  
indoleacetamide hy  
probable ClpA/B pr  
hypothetical prote  
ornithine decarbox  
ubiquitin carboxyl  
hypothetical prote

hypothetical protein T10B11.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T33517  
R:Minx, P.; Kemp, K.  
submitted to the EMBL Data Library, October 1998  
A:Description: The sequence of C. elegans cosmid T10B11.  
A:Reference number: Z21363  
A:Accession: T33517  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-549 <MIN>  
A:Cross-references: UNIPROT:Q9TZ11; EMBL:AF098993; PIDN:AAC67466.1; GSPDB:GN00019; CESP  
A:Experimental source: strain Bristol N2; clone T10B11  
C:Genetics:  
A:Gene: CESP:T10B11.2  
A:Map position: 1  
A:Introns: 26/1; 76/2; 109/3; 159/1; 187/3; 229/3; 353/1; 398/2; 434/1; 467/3

RESULT 1

T33517  
hypothetical protein T10B11.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T33517  
R:Minx, P.; Kemp, K.  
submitted to the EMBL Data Library, October 1998  
A:Description: The sequence of C. elegans cosmid T10B11.  
A:Reference number: Z21363  
A:Accession: T33517  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-549 <MIN>  
A:Cross-references: UNIPROT:Q9TZ11; EMBL:AF098993; PIDN:AAC67466.1; GSPDB:GN00019; CESP  
A:Experimental source: strain Bristol N2; clone T10B11  
C:Genetics:  
A:Gene: CESP:T10B11.2  
A:Map position: 1  
A:Introns: 26/1; 76/2; 109/3; 159/1; 187/3; 229/3; 353/1; 398/2; 434/1; 467/3

ALIGNMENTS

QY 523 SWNCDEVLHSP---AIEVRVHCOLVRLPAR 550  
||| ||| : : : : :  
Db 496 VWNLDGEILEQPKDEPLFKLHPQLISFGR 526  
||| ||| : : : : :  
RESULT 2  
T05162  
hypoetical protein F18E5.160 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: T05162  
R:Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Meyer, J.; et al. 1998  
submitted to the Protein Sequence Database, August 1998  
A:Reference number: Z15400  
A:Accession: T05162  
A:Molecule type: DNA  
A:Residues: 1-1240 <BEV>  
A:Cross-references: UNIPROT:O65419; EMBL:ALO22603  
A:Experimental source: cultivar Columbia; BAC clone F18E5  
C:Genetics:  
A:Map position: 4  
A:Introns: 44/3; 117/3; 178/3; 206/1; 237/2; 375/1; 416/1; 449/3; 474/1; 503/1; 529/1; 531/1  
A:Note: F18E5.160  
Query Match 13.8%; Score 417.5; DB 2; Length 1240;  
Best Local Similarity 28.9%; Pred. No. 1.3e-25;  
Matches 129; Conservative 75; Mismatches 162; Indels 81; Gaps 16;  
QY 140 WLQTLREMLEKLTSPKHLVFINPFGGKGQKRIYERKVPAPLFTLASITTDIIIVTEHAN 199  
||| ||| : : : : :  
Db 363 WCYKLRQYLDL-GRPKLLVFNVPFGGKSAREIFVKEVKLPEDADVQLEIQETKYQL 421  
||| ||| : : : : :  
QY 200 QAKETLYEINDKYDGI VCVGDMFSEVLHGLIGRTO-RSAGVDQNHPRAVLPSSLR 258  
||| ||| : : : : :  
Db 422 HAKEFVKSMVSKYDGI VCVGDMFSEVLHGLIGRTO-RSAGVDQNHPRAVLPSSLR 258  
||| ||| : : : : :  
QY 259 GIIPAGSTDCVYSTVGTSD---AETSAHLVVGDSLAMDVSSVHHNSTLLRSVSLG 314  
||| ||| : : : : :  
Db 469 GWNPAGTNGMIKSLDVTGLRCCANSATISIRGKRSVDVATIAQNTVF-FSVLMLA 527  
||| ||| : : : : :  
QY 315 YGFYGDIIKDSEKRWLGLARYDF-----SGLKTFLSHHCYEGTVSF 356  
||| ||| : : : : :  
Db 528 WGLIADIDIESEKPRWMSARIDFVCLVDFNYCIAVVKLALQRIICLLRYNGRI 587  
||| ||| : : : : :  
QY 357 LPAQ--HTVGPDRKPCACGFCVCRQSKQLEBEQKALYGLEAAEDVEEHWVQCKFL 414  
||| ||| : : : : :  
Db 588 LPAPGEGYQGPAS-----CSLYQEPHVSDEKVGQYQGPETKPEDLEWREMKGPV 637  
||| ||| : : : : :  
QY 415 AINATNMSCACRRSPRG-----LSPAHLGDSGLIIRKCSRNFNRLFLIRHTNQDQ 469  
||| ||| : : : : :  
Db 638 TIWLHN-----PWGSENTITAPAAKPSDGLDVLVKNCPLVLLS-LMRQTSSGTH 689  
||| ||| : : : : :  
QY 470 FDTFV-----EVVRVKKQFTSKHMEDESDLKEGGKRGFGHICSSHPSCCCTVSNSSW 524  
||| ||| : : : : :  
Db 690 VESPIVIVIKLTVKVKAFVLEPGALVDEPD--KEGIIDSDGEVLAR-----GKRTY 739  
||| ||| : : : : :  
QY 525 NCDGEVLHS-PAIEVRVHCOLVRLPAR 550  
||| ||| : : : : :  
Db 740 KCDQKALMSYDKLQVTVDQE--RLFCCK 764  
||| ||| : : : : :  
RESULT 3  
T38776  
hypoetical protein SPAC4A8.07c - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: T38776  
R:Skilton, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
submitted to the EMBL Data Library, August 1997  
A:Reference number: Z21751  
A:Accession: T38776  
A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-458 <SKE>  
A:Cross-references: UNIPROT:O14159; EMBL:Z98762; PIDN:CAB11477.1; GSPDB:GN00066; SPDB:SP1  
A:Experimental source: strain 972h-; cosmid c4A8  
C:Genetics:  
A:Gene: SPDB:SPAC4A8.07c  
A:Map position: 1  
A:Introns: 39/1; 101/1  
Query Match 9.6%; Score 291.5; DB 2; Length 458;  
Best Local Similarity 23.9%; Pred. No. 7e-16;  
Matches 105; Conservative 72; Mismatches 175; Indels 87; Gaps 16;  
QY 75 CSVPVSEIIAEEETDVHGKHQSGKQKMEKPYAFTVHCVKRARRHRWK-WAQVTFWCP 133  
||| ||| : : : : :  
Db 35 CSIPVS-----AKVDLNIIPFKNI-LWVDKTPNSVTLSYVSRSSKVATKCH--VDFVENS 87  
||| ||| : : : : :  
QY 134 EQLCHLWLQTLREMLEKLTSPKHLVFINPFGGKGQKRIYERKVPAPLFTLASITTDII 193  
||| ||| : : : : :  
Db 88 DQFC-----EVLDDVAYKIGKRSRRFIVFINPHGKGKAKHIWESEAEVPFSSAHSICEV 143  
||| ||| : : : : :  
QY 194 VTEHANOAKETLYEINDKYDGI VCVGDMFSEVLHGLIGRTO-RSAGVDQNHPRAVLP 253  
||| ||| : : : : :  
Db 144 LTRKDHAKSIANKLDVGSYDGLSVGGDGLFHEVINGL-----GERDDYLEAFKLP 195  
||| ||| : : : : :  
QY 254 SSLRIGIIPAGSTDCVYSTVGTSDAETSAHLVVGDSLAMDVSSVHHNSTLLRSVSL 313  
||| ||| : : : : :  
Db 196 ----VCMIPOGSGNAFSYNATGQKPALEILKGRPTSDLMTFEQKGK-KAYSFLTA 250  
||| ||| : : : : :  
QY 314 GYGFYGDIIKDSEKRWLGLAR-YDFSLKLTFLSHHCYEGTVSFPAQHTVGSPRDRKP- 371  
||| ||| : : : : :  
Db 251 NYGIIADCIDIGTENWRFMGENRAYLGFFLRLF-----OKPD 286  
||| ||| : : : : :  
QY 372 --CRAGFCVCRQSKQLEBE--BQKALYGLEAED-----VSEWQVVC 410  
||| ||| : : : : :  
Db 287 WKCSIENDVWSSDRTEIKHMYEKSKNLAPMSESSDKTSTVSTPESHLLTFEINDLSIFC 346  
||| ||| : : : : :  
QY 411 GKFLAINATNMSCACRRSPRGSLSPAHLGDSGLIIRKCSRNFNRLFLIRHTNQDQ 470  
||| ||| : : : : :  
Db 347 AGLLPYATPAK-----MFPAAANDGLIDVIVYS---KQFRKSLLSMTQLDNG 394  
||| ||| : : : : :  
QY 471 DFTF---VEVVRVKKQFT 486  
||| ||| : : : : :  
Db 395 GFYYSKHLNYYKRSFRPT 413  
||| ||| : : : : :  
RESULT 4  
S51398  
hypoetical protein YLR260w - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein L8479.7  
C:Species: Saccharomyces cerevisiae  
C:Date: 05-May-1995 #sequence\_revision 12-May-1995 #text\_change 09-Jul-2004  
C:Accession: S51398  
R:Miller, N.  
submitted to the EMBL Data Library, November 1994  
A:Description: The sequence of S. cerevisiae cosmid 8479.  
A:Reference number: S51395  
A:Accession: S51398  
A:Molecule type: DNA  
A:Residues: 1-687 <MIL>  
A:Cross-references: UNIPROT:Q06147; EMBL:U17244; NID:g577171; PIDN:AAB67377.1; PID:g5771.1  
C:Genetics:  
A:Gene: SGD:LCB5; MIPS:YLR260w  
A:Cross-references: SGD:S0004250  
A:Map position: 12R  
Query Match 8.2%; Score 248.5; DB 2; Length 687;  
Best Local Similarity 21.7%; Pred. No. 4.1e-12;  
Matches 99; Conservative 75; Mismatches 175; Indels 107; Gaps 14;  
QY 72 ADACSVPVSEIIAEEETDVHGKHQSGKQKQNM---EKPYAFTVHCV-KRARRHRKWAQV 127  
||| ||| : : : : :  
Db 174 SEKDLDELESSQKKERKGNLSRGSNSSSLLTSRSPFTKLVEVIFARPRHDVVPKRV 233  
||| ||| : : : : :

```
QY 128 TWCPEQLCHLWLOTLEWLEKL-----TSRPKHLVFINPGGKGQKRIYERKVA 180
Db 234 SLYIDYKPHSSHLKEEDDLVEILKRSYKNTFRNKSI FVIINPGGKGKAKLFWTKAK 293
QY 181 PLFTLASITTDIIVTEHANQAKETLYEINIDKYDGVICVGGDMFSEVLHGLIGRTQRSA 240
Db 294 PELLASRCSIEVYTKYPGHAIEAREMDIDKYDTIACASGDGPIHEVINGLYQR----- 348
QY 241 GVDQNHPRAVLPSSLRIGIIPAGSTDCVCYSTVGTGSDAETSALHIVVGDLSLAMDVSSH 300
Db 349 -----PDHVKAFFNIAITEIPCGSGNAMSVSCHWTNPNSTYTLCLIKSIETRIDLMCCS 402
QY 301 HNSTLLRY-SVSLIG--YGFYGDIIKQSEKKWGLIARYDF----- 338
Db 403 QFSYAREHPKJLSQTYGLIAETDINTEFIRWGMGPARELGVAFNIIQKKYPCIEIYK 462
QY 339 -----SGLKTFSLSHHCYEGTVSFLPAQHTVGSPPDRKPCRAGCF----- 377
Db 463 YAAKSKVELKNHYLEHKNKGSLEF---QH-ITWKNQDNEDCDNYYENEYETENEDEDEDA 518
QY 378 -----VCRSQKQLEEBOKKALYGLEAAEDV-EWQ-----VWCG 411
Db 519 DADDESHLISRDLDASSAQIKEEDFKIKYPLD--EGIPSDWERLDPNISNNLGIFTVG 576
QY 412 KFLATNATMNSCACRRSPRLGSPAHLGDGSSDLIL 447
Db 577 KMPYVAADTK-----FFPAALPDSGDTMDMVI 602

RESULT 5
T19707
hypothetical protein C34C6.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T19707
R:Percy, C.
submitted to the EMBL Data Library, October 1995
A:Reference number: Z19167
A:Accession: T19707
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-473 <WIL>
A:Cross-references: UNIPROT:Q18425; EMBL:Z66494; PIDN:CAA91259.1; GSPDB:GN000020; CESP:C34C6
A:Experimental source: clone C34C6
C:Genetics:
A:Gene: CESP:C34C6.5
A:Map position: 2
A:Introns: 82/1; 126/1; 158/3; 276/1; 311/3; 427/3

Query Match 8.1%; Score 245.5; DB 2; Length 473;
Best Local Similarity 24.3%; Pred. No. 4.3e-12;
Matches 109; Conservative 66; Mismatches 159; Indels 115; Gaps 18;

QY 149 EKLTS-----RPK-----HLLVFINPGGKGQKRIYERKVAFLFLASITTDIIVTE 196
Db 67 EOLTSVILSRKPPPEQCRGNLLVFINPNSGKSLFTFANTVGPDKLSIRYEVVVTT 126
QY 197 HANQAKETLY-EINIDKYDGVICVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSS 255
Db 127 GPNHARNVLMYKADLGKNGVLLSGDLGVFALNGILCRE-----AFRIPTT 175
QY 256 LRIGIIPAGSTD---CVCYSTVGTGSDAETSALHIVVGDLSLAMDVSSHNSTLLRYSVSL 312
Db 176 LPIGIVPSGSGNGLLCSVLKYGTQWNEKS----VMERALEIATSPTAKAESVALYSVK 231
QY 313 -----LGYGYGDIIDKSEK-KRWGLIARYDFSLKTFIASHHCYEGTVSFLPAQH 361
Db 232 DNQSVASFLSTGWLGLADIDIDSEKWRKSLGHHRFTVMGFI RSCNLSRYKGLTYRYPK- 290
QY 362 TVGSPDRKPCRAGCFVCRSQKQLEEBOKKALYG----- 396
Db 291 ----PKGFHPSNVFSVYEKTTQQRIDDSKVKTNGSVSDSEETMETKFNQWTLPPDSDET 346
```

```
QY 397 -LEAARDVEHQVCGKFLAINATMNSCACRRSPRLGSPAHLGDGSSDL--ILIRK-S 452
Db 347 LAVGSSDLEETVVIDNEFVNIYAVTLISHIAADGP--FAPSARKLEDNRHLSYILWKDIGT 404
QY 453 RNFNLFRLIRHNTQODQDFTFVEYVRVKKQFTSKHMEDESDLK---EGGKRFGHIC 509
Db 405 RVNIKAYLLA-IEHTHLDLPFV-----KHVEVSSMKLEVISEG-----SHVV 446
QY 510 SSHPSCCCTVSNSSWNCDEVLHSPAIEV 538
Db 447 -----LDGEVVDTKTIEV 459

RESULT 6
S67059
hypothetical protein YOR171c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein O3615
C:Species: Saccharomyces cerevisiae
C>Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
C:Accession: S67059
R:Bordonne, R.; Camasses, A.; Madania, A.; Martin, R.P.; Poch, O.; Tarasov, I.A.; Winsor
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67032
A:Accession: S67059
A:Molecule type: DNA
A:Residues: 1-624 <BOR>
A:Cross-references: UNIPROT:Q12246; EMBL:Z75078; NID:g1420415; PID:g1420417; GSPDB:GN000
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:LCB4; MIPS:YOR171C
A:Cross-references: SGD:S0005697
A:Map position: 15R

Query Match 7.0%; Score 210.5; DB 2; Length 624;
Best Local Similarity 26.3%; Pred. No. 4.7e-09;
Matches 66; Conservative 46; Mismatches 102; Indels 37; Gaps 6;

QY 143 TLREMLEKL---TSRPKHLVFINPGGKGQKRIYERKVAFLFLASITTDIIVTEHAN 199
Db 211 TVEILEKSYENSKRNSILVINPHGKGKTAKNLFITKARPILVESCKIEIAYTKYAR 270
QY 200 QAKETLYEINIDKYDGVICVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIG 259
Db 271 HAIDIAKDLSDIKYDTIACASGDGPIYEVINGLYRRPDR---VDAFNKLAVTQ----- 320
QY 260 IIPAGSTDVCYSTVGTGSDAETSALHIVVGDLSLAMDV---SSVHNSLTLRYSVSLG 316
Db 321 -LPCSGNAMSVISCHWTNPNSTYTLCLIKSIETRIDLMCCSQPSTMNWPKLSFLSQTYG 379
QY 317 FYGDIIDSEKKRWGLIARYDFSLKTFIASHHCYEGTVSFLPAQHTVGSPPDRKPCRAGC 376
Db 380 VIAESDINTEFIRWGPVRFNLG-----VAFNIQG-----KKYPCVFPV 419
QY 377 FVCRSQKQLE 387
Db 420 KYAAKSKKELK 430
```

```
RESULT 7
AG1665
hypothetical protein homolog lin1865 [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AG1665
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Varquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
```



Db 184 GK--LFQGEI-----MLFLVLTNSVGFEK-----LAPDSSLNDGNFDMILDK 225  
QY 450 KCSRNFNLFRLI-----RHTNQODQDFTFVEYVRK 481  
Db 226 KANLAEFIRVATMALRGEHINDQ---HIIYTKANRVK 259

RESULT 10  
AH1769  
conserved hypothetical protein lin2702 [imported] - Listeria innocua (strain Clip11262)  
C:Species: Listeria innocua  
C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AH1769  
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
A.; Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AH1769  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-306 <GLA>  
A:Cross-references: UNIPROT:Q927T6; GB:AL592022; PIDN:CAC97928.1; PID:g16415238; GSPDB:G  
A:Experimental source: strain Clip11262  
C:Genetics:  
A:Gene: lin2702

Query Match 4.5%; Score 136.5; DB 2; Length 306;  
Best Local Similarity 19.6%; Pred. No. 0.0021;  
Matches 62; Conservative 55; Mismatches 91; Indels 109; Gaps 14;  
QY 156 KHLVFINPFGKGKRI-----VERKVAPLFTLASITTDIIIVTEHANQAKE 203  
Db 3 KAMIIYNPAGKNKFKLLPDAEKILTNADFEVLVP-STPAPKSTTLI---AKQAAE 57

QY 204 TLYEINIDKYGIVCGDGMFSEVLHGLIGRTORSAGVDQNHPRVLPVSSLRIGIIPA 263  
Db 58 AGYEV-----VIAAGDGTNEVNGLMQVEKRP-----KLGIPLV 93

QY 264 GSTDCVCVSTGTSDAETSALHIVVGDLSLMDVSSVHHNSTLLRYS-----VSLLOYGYG 319  
Db 94 GTTNDYARALNPAKDFLEALQIIAQETIRVDIGKANETFEFINNAAGGRITEITYA--- 150

QY 320 DIIKDSEKKRWGLARYDFSGLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVC 379  
Db 151 --VKESMSKW-GRLAYLFSGL-----TVLP----- 173

QY 380 ROSKOOLBEEQKALYGLEAADVEWQVVCCKFLAINATN-----MSCACRRSPRGLSPA 435  
Db 174 ----KLSPVNVVEIYN---REIFKGEIL---LFFVNKTSVGGMETLC-----PP 213

QY 436 AHLGDGSSDLILIRKCS 452  
Db 214 AOLNSGMFELLILKKVS 230

RESULT 11  
T16422  
hypothetical protein F52C9.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C:Accession: T16422  
R;Favella, T.  
submitted to the EMBL Data Library, November 1995  
A:Description: The sequence of C. elegans cosmid F52C9.  
A:Reference number: Z18511  
A:Accession: T16422  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA

A:Residues: 1-732 <FAV>  
A:Cross-references: EMBL:U39850; NID:g1055052; PID:g1055058; PIDN:AAA81060.1; CESP:F52C9  
C:Genetics:  
A:Gene: CESP:F52C9.3  
A:Introns: 63/1; 106/3; 148/2; 303/1; 347/1; 408/2; 650/2

Query Match 4.5%; Score 136; DB 2; Length 732;  
Best Local Similarity 19.9%; Pred. No. 0.0075;  
Matches 76; Conservative 56; Mismatches 128; Indels 122; Gaps 18;  
QY 153 SRPKHLVFINPFGKGKRIYERKVAPLFTLASITTDIIIVTEHANQAKETLYEINIDK 212  
Db 65 TRPKRVFLVNVNVEGNSRCFDPQFNKNALPLFHLAGQVDVVKADNQAOLEALAGAVDTQE 124

QY 213 YDGIWCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRVLPVSSLRIGIIPAG----- 264  
Db 125 ADILVWVGDDGTIGTVTGTIFRNREK-----AQLPVGFYPGGYDNLWLKR 169

QY 265 -----STDCV---CYSTVGTSDAETSALHIVVGDLSLMDVSSVHHNSTLL--RYSVSL 313  
Db 170 MLPSVFNSDDVRHACETAMAVIEDQKSVY-----AFELTT--EGSTLAPYGLGDV 220

QY 314 GYGFVGDIIKDSEKKRW--LGLAR---YDFSGLK---TFLSHHCYEGTVSFLPAQHTVG 364  
Db 221 SAGHFRQ--IEDTRKKFWFSNAKRWAYFWEMLRGPAPIECHVEYET----- 268

QY 365 SPRDRKPCRAGCFVCROSQKOLEBEQKALYGLEAADVEWQVVCCKFLAINATNMSCA 424  
Db 269 -----C-AGCEKCR-PKPIIEAPQWR-----WWHVLTG----- 294

QY 425 CRRSPRGLSPAHLGDGSSDL--ILIRKCSRPNFL-----RFLRHTNQODQDFTFVEV 477  
Db 295 -----TPKYXNDGQKDYTGIIINEKCGKHELDTHGAEFLIENEQMSD----- 337

QY 478 YRVKKFOFTSKHMEDESDLKE 499  
Db 338 YSQIRFMGDDPYMPEEFENNE 359

RESULT 12  
AH1528  
conserved hypothetical protein lin0768 [imported] - Listeria innocua (strain Clip11262)  
C:Species: Listeria innocua  
C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AH1528  
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
A.; Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AH1528  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-309 <GLA>  
A:Cross-references: UNIPROT:Q92DPS; GB:AL592022; PIDN:CAC96000.1; PID:g16413219; GSPDB:G  
A:Experimental source: strain Clip11262  
C:Genetics:  
A:Gene: lin0768

Query Match 4.3%; Score 130.5; DB 2; Length 309;  
Best Local Similarity 22.3%; Pred. No. 0.0065;  
Matches 42; Conservative 33; Mismatches 88; Indels 25; Gaps 5;  
QY 159 LVFINPFGKGKRIYERKVAPLFTLASITTDIIIVTEHANQAKETLYEINIDKDVIGVC 218  
Db 5 LLIVNPSGKEGK-YYQKTEEVLRKRYDEVRLETKAGDATEFASWASEQGFDAVIA 63

QY 219 VGGDGMFSEVLHGLIGRTORSAGVDQNHPRVLPVSSLRIGIIPAGSDVCYSTVGTSFSD 278  
Db 64 MGGDGLNETINGL-----AIHEKRP-----DFGFIPLGTVNDLARSVGIFLKL 106





Db	162	YIKLEK-----LPOI-----SPTDVR-----	178
Qy	397	LEAAEDVEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDLILIRKCSRNF	456
Db	179	IEYDGKLFEGEIM--MFLVSN-TNSVGGFER----LAPNASLRDGMFDFIIVKKTSPPEF	231
Qy	457	LRFL-----IRHTNQDQDFTFEVEYRVK	481
Db	232	LHLAGLALRGEHIKHPK-----VLYVQANRIK	258

Search completed: September 3, 2005, 04:38:45  
Job time : 25.8463 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2005, 19:28:44 ; Search time 317.875 Seconds  
(without alignments)  
2892.921 Million cell updates/sec

Title: US-10-631-958-11  
Perfect score: 3025  
Sequence: 1 HEANGPAPLGVAPPWRT.....QLVRLPARGIENPKPDHS 562

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO spool/US10631958/runat 02092005 165811 4432/app query.fasta\_1.1941  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPTCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10631958 @cgn 1.1 397 @runat 02092005 165811 4432 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/5A COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/5B COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTCUT COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2931.5	96.9	4432	4	US-09-774-528-148 Sequence 148, App
2	813	26.9	2064	4	US-09-270-767-14306 Sequence 14306, A
3	585.5	19.4	1084	4	US-09-270-767-15155 Sequence 15155, A
4	302.5	10.0	901	4	US-09-270-767-30448 Sequence 30448, A
5	288.5	9.5	1857	4	US-09-970-516-3 Sequence 3, Appli
6	288.5	9.5	2380	4	US-09-817-676A-13 Sequence 13, Appl
7	280.5	9.3	1783	4	US-09-949-016-1155 Sequence 1155, Ap
8	269	8.9	2698	4	US-09-817-676A-11 Sequence 11, Appl
9	266.5	8.8	1205	4	US-09-959-897-1 Sequence 1, Appli
10	264.5	8.7	1155	4	US-09-970-516-1 Sequence 1, Appli
11	257.5	8.5	1533	4	US-09-205-258-90 Sequence 90, Appl
12	247.5	8.2	1149	4	US-09-970-516-5 Sequence 5, Appli

13	228	7.5	1611	4	US-09-248-796A-1756 Sequence 1756, Ap
14	217.5	7.2	1764	4	US-09-949-016-3940 Sequence 3940, Ap
15	210.5	7.0	1875	4	US-09-614-221A-399 Sequence 399, App
16	199.5	6.6	2462	4	US-09-620-312D-796 Sequence 796, App
17	189.5	6.3	1012	4	US-09-270-767-12677 Sequence 12677, A
18	153	5.1	942	4	US-09-107-532A-3500 Sequence 3500, Ap
19	153	5.1	1239	4	US-09-603-208A-263 Sequence 263, App
20	144	4.8	959	4	US-09-270-767-8490 Sequence 8490, Ap
21	144	4.8	959	4	US-09-270-767-21772 Sequence 21772, A
22	140	4.6	7195	4	US-09-949-016-12897 Sequence 12897, A
23	140	4.6	7198	4	US-09-949-016-15682 Sequence 15682, A
24	132.5	4.4	4411529	3	US-09-103-840A-1 Sequence 1, Appli
25	129	4.3	3120	4	US-09-902-540-5505 Sequence 5505, Ap
26	129	4.3	3122	4	US-09-902-540-469 Sequence 469, App
27	126	4.2	10717	4	US-09-902-540-991 Sequence 991, App
28	125	4.1	485	4	US-09-270-767-31476 Sequence 31476, A
29	123.5	4.1	915	4	US-09-134-000C-1961 Sequence 1961, Ap
30	119.5	4.0	3089	4	US-09-634-238-188 Sequence 188, App
31	119	3.9	17612	3	US-08-911-853-29 Sequence 29, Appl
32	119	3.9	17612	3	US-09-479-409-29 Sequence 29, Appl
33	119	3.9	17612	3	US-09-479-453-29 Sequence 29, Appl
34	119	3.9	4403765	3	US-09-103-840A-2 Sequence 2, Appli
35	118.5	3.9	11575	4	US-09-938-956-5 Sequence 5, Appli
36	118	3.9	912	4	US-09-107-532A-1479 Sequence 1479, Ap
37	118	3.9	8971	4	US-09-949-016-15668 Sequence 15668, A
38	117	3.9	2963	3	US-09-232-200-60 Sequence 60, Appl
39	117	3.9	2963	3	US-09-232-197-60 Sequence 60, Appl
40	117	3.9	2963	3	US-09-232-201-60 Sequence 60, Appl
41	117	3.9	2963	4	US-09-232-195-60 Sequence 1079, Ap
42	116.5	3.9	885	4	US-09-583-110-1079 Sequence 1079, Ap
43	116.5	3.9	948	4	US-09-107-433-518 Sequence 518, App
44	115	3.8	2469	4	US-09-252-991A-5908 Sequence 5908, Ap
45	115	3.8	3135	4	US-09-252-991A-5922 Sequence 5922, Ap

ALIGNMENTS

RESULT 1  
US-09-774-528-148  
; Sequence 148, Application US/09774528  
; Patent No. 6743619  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Zhou, Ping  
; APPLICANT: Goodrich, Ryle  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Zhang, Jie  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6743619el Nucleic Acids and  
; FILE REFERENCE: Polypeptides  
; CURRENT APPLICATION NUMBER: US/09/774,528  
; NUMBER OF SEQ ID NOS: 441  
; SOFTWARE: pt\_FL\_genes Version 2.0  
; SEQ ID NO 148  
; LENGTH: 4432  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1129) .. (2817)  
US-09-774-528-148  
Alignment Scores:

Pred. No.:	6.52e-279	Length:	4432
Score:	2931.50	Matches:	548
Percent Similarity:	98.21%	Conservative:	2
Best Local Similarity:	97.86%	Mismatches:	7
Query Match:	96.91%	Indels:	3
DB:	4	Gaps:	1
US-10-631-958-11 (1-562) x US-09-774-528-148 (1-4432)			
Qy	6	GlyProAlaProLeuGly-----ValArgAlaProProAlaTrpArgThrSerPro	22
Db	1135	GGTCCCGCCCTATAGACAAGACAGTCAACAGGAGCTGGCAGGCGCTGGCGGACGAGCCG	1194
Qy	23	AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLys	42
Db	1195	CGCGCGGAGATGGGGCGACGGGGCGCGGAGCGCGTGCATTCGTCGCTGTGGGTGAAG	1254
Qy	43	GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSer	62
Db	1255	CAGCAGCGCTGCGCGGTAGCGCTGGAGCCCGCGGGCTCTGCTGCGCTGTGGCGGAGC	1314
Qy	63	ProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIle	82
Db	1315	CCGGGCGCGGAGCGCGGCGCCCGCGCGGATGCTCTGTCGCTGTATCTGAGATC	1374
Qy	83	IleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLys	102
Db	1375	ATCCCGCTTGAGGAAACAGAGCTTTCACGGGAAACATCAAGGCAGTGGGAAATGGCAGAA	1434
Qy	103	MetGluLysProTrpAlaPheThrValHisCysValLysArgAlaArgHisArgTrp	122
Db	1435	ATGGAAAGCCCTACGCTTTTACAGTTTACCTGTGTAAAGAGACACGCGCACCGCTGG	1494
Qy	123	LysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGln	142
Db	1495	AAGTGGCGCAGGTGACTTTCTGTGTTCAGAGGAGCAGCTGTCTCACTTGTGGCTCGAG	1554
Qy	143	ThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIle	162
Db	1555	ACCTCTGGGAGATGCTGGGAAGCTCAGCTCCAGACCAAGCATTTTACTGGTATTTATC	1614
Qy	163	AsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrgluArgLysValAlaProLeu	182
Db	1615	AACCCGTTGGAGGAAAGGACAGGACAGCGGATATATGAAGAAAGTGGCCACCACTG	1674
Qy	183	PheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLys	202
Db	1675	TTCACTTAGCTCCATCACCACCTGACATCATCTGTTACTGAACATGCTAATCAGGCCAAG	1734
Qy	203	GluThrLeuTyrgluIleAsnIleAspLysTyrgaspGlyIleValCysValGlyGlyAsp	222
Db	1735	GAGACTCTGTATGAGATTAAACATAGACAAATACACGGCATTCGTCCTGTGTGGCGGAGAT	1794
Qy	223	GlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyVal	242
Db	1795	GGTATGTTACGAGAGTGCTGCACGCTGTGATTCGGAGGACCGACAGAGCGCCCGGGTC	1854
Qy	243	AspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIlePro	262
Db	1855	GACCAGAAACACCCCGCGGTGTCTGCTGCCAGTAGTACGCTCCGGAATGGAATCATTTCCC	1914
Qy	263	AlaGlySerThrAspCysValCysTyrgSerThrValGlyThrSerAspAlaGluThrSer	282
Db	1915	GCAGGGTCAACGGCATCGGTGTGTATCTCAACCGTGGGACCAACGCGACGAGAAACCTCG	1974
Qy	283	AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerValHisHisAsn	302
Db	1975	CGCTGCATATCTGTTGTGGGACTCGTGGCCATGGATGTGTCTCACTCCACCAAC	2034
Qy	303	SerThrLeuLeuArgTrpSerValSerLeuLeuGlyTyrglyPheTyrglyAspIle	322
Db	2035	AGCACCTCTTCGCTACTCGTCTCGTCTCGTGGGCTACGCTTCTACGGGACATCATC	2094

Qy	323	LysAspSerGlnIlysLysAsrGTrpLeuGlyLeuAlaArgTyrAppSheSerGlyLeuLys	342
Db	2095	AAGACACAGTCAGAGAAGAACCGTGCTGGGCTCTGCCAGATACAGACTTTTCAGGTTTAAAG	2154
Qy	343	ThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr	362
Db	2155	ACCTTCCTCTCCACCACTGCTATGGAGGACAGTGTCCTTCCTCCCTGCACAAACACACG	2214
Qy	363	ValGlySerProArgAspArgLysPysProCysArgAlaGlyCysPheValCysArgGlnSer	382
Db	2215	GTGGGATCTCCAGGGATAGAAAGCCCTGCGGGCAGGATCGTTTGTTCGAGCAAAAGC	2274
Qy	383	LysGlnGlnLeuGluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp	402
Db	2275	AAGCAGCAGCTGGAGGAGGACGAGAAAGCACTGTATGTTTGGAAAGCTGCGGAGAC	2334
Qy	403	ValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSer	422
Db	2335	GTGGAGGAGTGGCAAGTCGTCGTGGGAAGTTTCTGSCCATCAATGCCACAAACATGTCC	2394
Qy	423	CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer	442
Db	2395	TGTCCTGTGTCGCGAGCCCGAGGGGCTCTCCCGGCTGCCACTTTGGGAGACGGGTCT	2454
Qy	443	SerAspLeulleLeulleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArg	462
Db	2455	TCTGACCTCATCTCATCCGGAAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGG	2514
Qy	463	HisThrAsnGlnInAspGlnPheAspPheThrPheValGluValTyrArgValLysLys	482
Db	2515	CACACCAACAGCAGGACCAAGTTTGACTTCACATTTTGTGAAGTTTATCGCGTCAAGAAA	2574
Qy	483	PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLys	502
Db	2575	TTCAGATTACGTCGAAGCACATGGAGGATGAGGACGCACTCAAGGAGGGGGGGAAG	2634
Qy	503	LysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSer	522
Db	2635	AAGCGCTTTGGGCACATTTGCGAGCAGCACCCCTCTCTGCTGCACCGTCTCCACAGC	2694
Qy	523	SerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCys	542
Db	2695	TCCTGGAACTGCGACGGGGAGGTCCTGCACAGCCCTGCCATCGAGGTGAGATCCACTGC	2754
Qy	543	GlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer	562
Db	2755	CAGCTGGTTTCGACTCTTTGTCACGAGGAATTGAAGAGAAATCCGAAGCCAGACTCACACAGC	2814
RESULT 2			
US-09-270-767-14306			
; Sequence 14306, Application US/09270767			
; Patent No. 6703491			
; GENERAL INFORMATION:			
; APPLICANT: Homburger et al.			
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster			
; FILE REFERENCE: File Reference: 7326-094			
; CURRENT APPLICATION NUMBER: US/09/270,767			
; CURRENT FILING DATE: 1999-03-17			
; NUMBER OF SEQ ID NOS: 62517			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 14306			
; LENGTH: 2064			
; TYPE: DNA			
; ORGANISM: Drosophila melanogaster			
US-09-270-767-14306			
Alignment Scores:			
Pred. No.:		4,55e-70	Length: 2064
Score:		813.00	Matches: 207
Percent Similarity:		46.33%	Conservative: 102
Best Local Similarity:		31.03%	Mismatches: 218
Query Match:		26.88%	Indels: 141
DB:		4	Gaps: 17



```

; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15155
; LENGTH: 1084
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-15155

Alignment Scores:
Pred. No.:      4,97e-48      Length:      1084
Score:          585.50      Matches:      125
Percent Similarity: 51.14%      Conservative: 55
Best Local Similarity: 35.51%      Mismatches: 103
Query Match:      19.36%      Indels:      69
DB:              4          Gaps:      6

US-10-631-958-11 (1-562) x US-09-270-767-15155 (1-1084)

Qy 154 ArgProLysHisLeuValPheIleAsnProPheGlyGlyLysGlnGlyLysArg 173
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 1044 CGAGTCGCGCGCTTTTGGTCTTTTATAAACCCCTATGGAGGTGCGAAGGGCGCTCAG 985

Qy 174 IleTyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIle 193
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 984 ACCTATGAGCGCCATGTGAGACCCTATTTTCCAGGCTTCGCGCGTAGACGCCACGTGTATC 925

Qy 194 ValThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyr 213
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 924 ACCACTCAGAGGGCAACCAAGTGAAGACATCTTCTGAGCCATGATCTGGGAGTATAC 865

Qy 214 AspGlyIleValCysValGlyLysAspGlyMetPheSerGluValLeuHisGlyLeuIle 233
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 864 GATCGGTTTGTGTGTCGAGGGCGATGGCACCGTACGAGAGTTCATCAACGCGACTGATA 805

Qy 234 GlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValPro 253
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 804 TTCGGTCAAAATCGAGAGTTGGGACTGGACGAAACAGCGGCCACCATCATTTCCAAGACCG 745

Qy 254 SerSerLeuArgIleGlyIleProAlaGlySerThrAspCysValCysTyrSerThr 273
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 744 ---GCTCTCCAGTGGGTGTATCCCGCTGGAGACCCGACACCATTGGGTATAGTATG 688

Qy 274 ValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAla 293
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 687 CACGGCAGCGCGGATGTGAGGACAGCGGCTATCCATGTGATTCTGGGCCACGATCGGGGA 628

Qy 294 MetAspValSerSerValHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeu 313
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 627 TTGGATGTGTGAGTGTGAGCAATGGCCAGTCCCTGCTCAGATTCTGTGCAGTGTCTCG 568

Qy 314 GlyTyrGlyPheTyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyVal 333
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 567 AGCTACGGGTACTCTGGCGCATGTGGCACCCGACAGCGGAGAACTACCGCTGGATGGGACCG 508

Qy 334 AlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThr 353
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 507 CGCCGGTACAGTACAGTGGGCGTCAAGGCCCTTCCTGAATAATCGCGGCTATGACGCCGAA 448

Qy 354 ValSerPheLeu-----ProAlaGlnHisThrValGly 364
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 447 CTGAGAAATGTTAAGAGCCCGATCTTCTACTGACACGCGCGTGGAGGACATTCGCGAG 388

Qy 365 SerProArgAsp-----ArgLysProCysArgAlaGlyCys 376
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 387 AGTCCGATAGTGTGTGCTCGTGGGAGTCAGTGCCATCCCTGCTGCTATGCCAATTGC 328

Qy 377 PheValCys-----ArgGlnSerLysGlnGlnLeuGluGlu 388
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 327 CAGCGCTGACGCTTCGCCAGCAGCATACAGGAACAGCGATCCTCATTTGTTTCATCCAAAG 268
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

```

```
Db 350 ATGAGAACCTCGGCAACATCAGTGAAGGTTGTCGGGGCAATTTCTTATGATCGC 409
Qy 418 AlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHis 437
Db 410 GCGCGAAACATAACCTGCGCTGCGCAGGAGTCCCAATGGCATCTCCGTTTACAGTCAT 469
Qy 438 LeuGlyAspGlySerSerAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 457
Db 470 CTGGGTGATGGTTGCTGACCTGATCTCTGTAAGAAAGACCTCACTGCTCAACAACTG 529
Qy 458 ArgPheLeuLeuArgHisThrAsnGlnGln---AspGlnPheAspPheThrPheValGlu 476
Db 530 CGTTTCTGCTCAACACGCGGGCAGAGTGTGATATTCGCAATTTGCTTTTGTAGAG 589
Qy 477 ValTyrArgValLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAsp 496
Db 590 GTATATCGCACAGGAAATTCGGTTTCAGAACTTTTTCGCCAGCGAGGAGTACAGC 649
Qy 497 LeuLysGluGlyGlyValLysArgPheGlyHisLeuCys----- 509
Db 650 TTGCGAGGCTCT-----TGTCAGCCGATTACACCGCCTGAA 685
Qy 510 -----SerSerHisProSerCysCysThrValSerAsnSerSerTyrAsnCysAsp 527
Db 686 GAATATGACCGCCCATCTCTCC-----TCGACGGAGTTTTCAGCTGGAACGTGAT 736
Qy 528 GlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuValArgLeu 547
Db 737 CGAGAGGTGGTGACCGACTGGACATACCATGCGATCGCAATTGTCACTCATCGAGTGC 796
Qy 548 PheAlaArgGlyLeuGluGluAsnProLysPro 558
Db 797 TTCATGCGGGTCCCAATTCCTTACAGCAAGCCC 829
```

## RESULT 5

```
US-09-970-516-3
; Sequence 3, Application US/09970516
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; TITLE OF INVENTION: polynucleotides encoding sphingosine kinases
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1857
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1857)
; OTHER INFORMATION:
US-09-970-516-3
```

```
Alignment Scores:
Pred. No.: 2,648-18 Length: 1857
Score: 288.50 Matches: 97
Percent Similarity: 42.74% Conservative: 59
Best Local Similarity: 26.58% Mismatches: 173
Query Match: 9.54% Indels: 37
DB: 4 Gaps: 7
```

US-10-631-958-11 (1-562) x US-09-970-516-3 (1-1857)

```
Qy 18 TrpArgThrSerProAlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSer 37
Db 23 TGGCTGCAGCACCCCGCTCTCCATGGCGAGTTGGCTCTTACCAGCCCGAGGCCAC 82
Qy 38 ValLeuTrpValLysGlnGlnArg-----CysAlaValSerLeu 50
```

```
Db 83 GCTTTGCCCTCACCTTACATCGAGGCCCTGCAATACAGCGGCTGCGCCCAACCTG 142
Qy 51 GluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyProGlyAlaGlyAlaPro 70
Db 143 AAGCAGGCCCGGGG---GTGGCCTGGTCCCGTGGCCGAGGTTCTCAGGCTGCTGACCC 199
Qy 71 GlyAlaAlaPalaCysSerValProValSerGluLeuLeuLeuLeuLeuLeuLeuLeuLeu 90
Db 200 ---TCCGAAGACGCGACCCCTCAGACTCAGCGGCTACTTCTGCACTTAC-ACCTACCC 255
Qy 91 HisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLysProTyrAlaPheThr 110
Db 256 CGGGCGCGCGCGGGCGCGCGAGCCACTCGACCTTCCGGGCGAGATGGGGCCGCC 315
Qy 111 ValHisCysValLysArgAlaArgHisArgTrpLysTrpAlaGlnValThrPheTrp 130
Db 316 ACCTACGAAGAGAACCGTGGCGAGCCCGAGCGCTGGGCCACTGCGCCCTCACCTGTCTG 372
Qy 131 CysProGluGluGlnLeuLysHisLeuTrpLeuGlnThrLeuArgGluMetLeuGluLys 150
Db 373 -----CTCCGAGGACTGCCCTCGCGGATGGGAGATCACCCCTGAC 417
Qy 151 LeuThrSerArgProLysHisLeuLeuValPheIleAsnProPheGlyGlyGlyGln 170
Db 418 CTGCTACTCGCGCGCGCGCGGTTGCTTCTATTGTCATTCCTTTGGGGTTCGGGGCGTG 477
Qy 171 GlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThr 190
Db 478 GCCTGGCAGTGGTGTAGAACCACTGCTCTTCCCAGATCTCTGAAGCTGGGCTGCTCTTC 537
Qy 191 AspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIle 210
Db 538 AACCTCATCCAGACAGAACACAGACACACACCGCGGAGCTGGTCCAGGGGCTGAGCCTG 597
Qy 211 AspLysTyrAspGlyIleValCysValGlyCysGlyMetPheSerGluValLeuHis 230
Db 598 AGTGAGTGGGATGGCATCGTCACGCTCGGGAGACGGGCTGCCATGAGGTGCTGAAAC 657
Qy 231 GlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaVal 250
Db 658 GGGCTCTAGATCGC-----CCTGACTGGGAGGAGCTGTG 693
Qy 251 LeuValProSerSerLeuArgIleGlyIleProAlaGlySerThrAspCysValCys 270
Db 694 AAGATGCGCT-----GTGGGCATCTCCCTCGCGCTCGGGCAACCGCTGGCC 741
Qy 271 TyrSer-----ThrValGlyThrSerAspAlaGluThr 281
Db 742 GGACGACTGAACACGACGCGGGGATTTGAGCCGCTGGGCGCTCGACCTGTTGCTCAAC 801
Qy 282 SerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHis 301
Db 802 TGCTCACTGTTGTGTCGCGGGGTGGTGGCCACCCACTGACCTGCTCTCGTGACCGCTG 861
Qy 302 AsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIle 321
Db 862 GCCTCGGGCTCCCGCTGTTCTCTCTGCTGTCGCGCTGGGGCTTCGTGTCAAGATGTG 921
Qy 322 IleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeu 341
Db 922 GATATCCAGAGCGAGCGCTTCAGGGCGCTTGGGCGAGTGCCTGCTTTCACACTGGGACCG 981
Qy 342 LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis 361
Db 982 CTGGGCGCTCGCACACCTGACACCTACCGGAGCGGCTCTCTTACCTCCCGCCACTGTG 1041
Qy 362 ThrValGlySerPro 366
Db 1042 GAACCTGCTCGCCC 1056
RESULT 6
US-09-817-676A-13
```



Db 424 CTGCTACTCGCGCGCCCGGTTGCTTCTATTGGTCAATCCCTTTGGGGGTGCGGGCGCTG 483  
 Qy 171 GlyLysArgfileTyrgluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThr 190  
 Db 484 GCCTGGCAGTGGTGAAGAACACCAAGCTTCCCATGATCTCTGAAGCTGGGCTGTCTTC 543  
 Qy 191 AspileValThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrgluIleAsnIle 210  
 Db 544 AACCTCATCCAGACAGAACACGACAGAACCCGCGGAGCTGTCTCCAGGGGCTGAGCGCTG 603  
 Qy 211 AspLysTyrglyAspGlyIleValCysValGlyGlyAspGlyMetPheSerGluValLeuHis 230  
 Db 604 AGTGAGTGGGATGGCATCGTCACGGTCTCGGAGACGGGCTGCTCCATGAGGTGCTGAAC 663  
 Qy 231 GlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaVal 250  
 Db 664 GGGCTCTAGATCGC-----CTGACTCGGAGGAAGCTGTG 699  
 Qy 251 LeuValProSerSerLeuArgIleGlyIleProAlaGlySerThrAspCysValCys 270  
 Db 700 AAGATGCCT-----GTGGGCATCTCCCTCGCGCTCGGGCAACGCGCTGGCC 747  
 Qy 271 TyrSer-----ThrValGlyThrSerAspAlaGluThr 281  
 Db 748 GSGAGCATGAACACGACGCGGGGATTGTAGCCAGCGCTGGCGCTCGACCTGTTGCTCAAC 807  
 Qy 282 SerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerValHisHis 301  
 Db 808 TGCTCACTGTGTGTGTCGGGGTGTGGGCCACCCACTGGACCTGCTCTCGTGACGCTG 867  
 Qy 302 AsnSerThrLeuLeuArgTyrglySerValSerLeuLeuGlyTyrglyPheTyrglyAspIle 321  
 Db 868 GCCTCGGGCTCCCGCTGTTTCTCCTCTGCTGTGGCCCTGGGGCTGCTGTGCAGATGTG 927  
 Qy 322 IleLysAspSerGluLysLysArgTTPLeuGlyLeuAlaArgTyrglyAspPheSerGlyLeu 341  
 Db 928 GATATCCAGACGAGCGCTTCAGGGCGCTGGGCAGTCCCGCTCACACTGGGCACGCGTG 987  
 Qy 342 LysThrPheLeuSerHisCysTyrglyGluGlyThrValSerPheLeuProAlaGlnHis 361  
 Db 988 CTGGGCTCGCCACACTGCACACTACCGCGGAGCGCTCTCTTACCTCCCGGCACTGTG 1047  
 Qy 362 ThrValGlySerPro 366  
 Db 1048 GAACCTGCCTCGGCC 1062  
 RESULT 7  
 US-09-949-016-1155  
 ; Sequence 1155, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SEQ ID NO 1155  
 ; FASTSEQ for Windows Version 4.0  
 ; SEQ ID NO 1155  
 ; LENGTH: 1783  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-09-949-016-1155  
 Alignment Scores: 1.53e-17 Length: 1783  
 Pred. No.:

Score: 280.50 Matches: 143  
Percent Similarity: 37.82% Conservatives: 82  
Best Local Similarity: 24.03% Mismatches: 223  
Query Match: 9.27% Indels: 148  
DB: 4 Gaps: 27

US-10-631-958-11 (1-562) x US-09-949-016-1155 (1-1783)

QY 2 GluAlaAlaAsnGlyProAlaProLeu-----GlyValArgAlaProProAlaTrpArg 19  
DB 2 GAGGAGCGAGCGCGGAGTCCGCTCCAGCGGGCGCTCCAGTCCCTCAGACGTGGGCT 61  
QY 20 -----ThrSerProAlaAlaGluMetGlyAlaThrGlyAlaAlaGlu 33  
DB 62 GAGCTTGGGACGAGCTGGCTCCGCCGCCAGGCACTGTAGGGAACGGGGTGGCCCTCC 121  
QY 34 ProLeuGlnSerValLeuTrpValLys-----GlnGlnArgCysAlaValSerLeuGlu 51  
DB 122 CCAGCAAAACCGGACCGAGCTGGGTCCAGCGCGGAGGAATGACACCGGTGTCTCTACAG 181  
QY 52 ProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyProGlyAlaGlyAlaProGly 71  
DB 182 CCA-----CGGCT-CGGGGC---GGGGAAGCGAGCCCCAC 213  
QY 72 AlaAspAlaCysSerValProValSerGluIleIleAlaValGluGluThrAspValHis 91  
DB 214 AGCGGCGCTGGAGCGCGCTGGGC-----AGCACCGAT----- 249  
QY 92 GlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLysProTyrAlaPheThrVal 111  
DB 250 ---AGGAGCTCAAGCGAGGAGCGCCGCCAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 291  
QY 112 HisCysValLysArgAlaArgHisArgTrpLysTrpAlaGlnValThrPheTrpCys 131  
DB 292 -----ACAGCGCCAGGAGCCCGCTGGCAG 315  
QY 132 ProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArgGluMetLeuGluLys--- 150  
DB 316 CGGGAGCGCGGGTC-----GAGTTATGATCCAGCG 348  
QY 151 -----LeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 165  
DB 349 GCGGCGCGCGGGGGTCTCCGCGGCGCTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 408  
QY 166 GlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 185  
DB 409 GCGGCAAGGCAAGCGCTTCAGCTCTCCGAGTCACTGCGAGCGGCGGCGGCGGCGGCGGCG 468  
QY 186 AlaSerIleThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 205  
DB 469 GCTGAATCTCTTCCACGCTGATGCTCACTGAGCGGCGGAACCCAGCGCGGAGCTGGT 528  
QY 206 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyAspGlyMetPhe 225  
DB 529 CGGTGGGAGGAGCTGGCGGCTGGAGCGCTCTGGTGTCTGTCATGCTGGAGAGCGGCTGATG 588  
QY 226 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspLysAsn 245  
DB 589 CACGAGGTGGTGAACGGGCTCATGAGCGGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCG 645  
QY 246 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAlaGlySer 265  
DB 646 -----CCCTGTGTAGCTCTCCACGAGGCTCT 672  
QY 266 ThrAspCysValCys-----TyrSerThrValGlyThrSer 277  
DB 673 GCAACGCGCTGGAGCTTCTTGAAACCATATGCTGGCTATGACGAGTCAACCATGAA 732  
QY 278 AspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAla---MetAspVal 296  
DB 733 GACCTCTGACCACTGACGAGTATGTGTGCGCGGCTGCTGTACCATGACCTG 792  
QY 297 SerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGly 316

Db 793 CTGTCTGTGCACACGGCTTGGGGCTGGCGCTCTTCTGTGTCTCAGCTGGCTGGGCG 852  
QY 317 PheTyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyr 336  
DB 853 TTCAATCTGATGTGGACCTAGAGAGTATCGGCGTCTGGGAGAGATGCGCTTC 912  
QY 337 AspPheSerGlyLeuLysThrPheLeuSerHisCysTyrGluGlyThrValSerPhe 356  
DB 913 ACTCTGGGCACTTCTCGCTGTGGCGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTTAC 972  
QY 357 LeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCys 376  
DB 973 CTCCCTGTAGGAAGA---GTGGGTTCAGACACCTGCTCTCCCTC----- 1014  
QY 377 PheValCysArgGlnSerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGly 396  
DB 1015 GTTGTGTCTCCAGCAGCGGCGGCTAGATGCACAC-----CTTGTGCCA 1056  
QY 397 LeuGluAlaAlaGluAspVal---GluGluTrpGlnValValCysGlyLys----- 412  
DB 1057 CTGGAG-----GAGCAGTGCCTCTCTCTGGAATGTGGTGGCGGAGGACTTGTG 1110  
QY 413 ---PheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGly 431  
DB 1111 CTAACTCTGGCAGCTGTGCACCTGCACCTGGGAGTGTGTGTGTGTGTGTGTGTGTG 1170  
QY 432 LeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIleLeuLys--- 450  
DB 1171 CGCTGTGCAGCT-----GGGTCTCATGATCTGTCTCAGTGGCGGCGGGA 1215  
QY 451 CysSerArgPheAsnPheLeuArgPheLeuIle-----ArgHisThrAsn 465  
DB 1216 GTGTCTGTGTCATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1272  
QY 466 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 485  
DB 1273 -----GAGTATGATGCTCTCTGATATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
QY 486 ThrSerLysHisMetGluAspGlySerAspLeuLysGluGlyLysLysArgPhe 505  
DB 1321 -----TTGGAGCCCAAGAT-----GGAAAGGTGTGTGT 1350  
QY 506 GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn 525  
DB 1351 GCA----- 1353  
QY 526 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 545  
DB 1354 GTGATGGGAATTTGATGTGTAGCGAGCGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1413  
QY 546 ArgLeuPheAlaArgGlyIleGluGluAsnPro-----LysPro 558  
DB 1414 TGGATGTGATGCTGT 1458

RESULT 8  
US-09-817-676A-11  
; Sequence 11, Application US/09817676A  
; Patent No. 6800470  
; GENERAL INFORMATION:  
; APPLICANT: Spiegel, Sarah  
; APPLICANT: Kohama, Takafumi  
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,  
; TITLE OF INVENTION: Expression and Methods of Use Thereof  
; FILE REFERENCE: 00170/HG  
; CURRENT APPLICATION NUMBER: US/09/817,676A  
; PRIOR FILING DATE: 2001-03-26  
; PRIOR FILING DATE: 2000-04-03  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 2698

```
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (387)..(2237)
; PUBLICATION INFORMATION:
; TITLE: Molecular cloning and functional characterization of a
; TITLE: novel mammalian sphingosine kinase type 2 isoform
; JOURNAL: J. Biol. Chem.
; VOLUME: 275
; ISSUE: 26
; PAGES: 19513-19520
; DATE: 2000-06-30
; DATABASE ACCESSION NUMBER: AF245448
; DATABASE ENTRY DATE: 2000-06-27
US-09-817-676A-11

Alignment Scores:
Pred. No.: 4.04e-16 Length: 2698
Score: 269.00 Matches: 98
Percent Similarity: 41.64% Conservative: 59
Best Local Similarity: 25.99% Mismatches: 159
Query Match: 8.89% Indels: 62
DB: 4 Gaps: 8

US-10-631-958-11 (1-562) x US-09-817-676A-11 (1-2698)

Qy 44 GlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTTrpArgSerPro 63
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
427 CAATCTGTCAGCGGAGTTTGGTCTACCGCGCCACAGCGCCACGGTTGGCCCTACCC 486

Qy 64 Gly-----ProGlyAlaGlyAlaProGlyAlaAsp 73
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
487 TCACACACAGCCCTACACATACAGCGACTACGCCCAAGCCAGACCGCCCGAG 546

Qy 74 AlaCysSerValProValSerGluIlelle-----AlaVal 85
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
547 ATGGTCTA-GTCTCTCTGGATGAGTCTCGGGCTGTGGCACCTTGCAGAGCCGTAGCCCC 605

Qy 86 GluGluThrAspValHis-----GlyLysHisGlnGlySer 97
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
606 GAGGACACTCAGCGCTTCTCGCATCTACCTACCCACGTCGCGCGTGGACCCCGAGCGGG 665

Qy 98 GlyLysTrpGlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArgAla 117
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
666 CGCAGAGCTACGCGACCTTCGCGCGCATGGGGCCACCACTTATGAGGAGATCGTGCA 725

Qy 118 ArgArgHisArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCys 137
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
726 GAGGCCAGCGCTGGGCCCACTGCGCTC-----CTCCGA 767

Qy 138 HisLeuTrpLeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHis 157
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
768 GGAGTGCTCTGTGAGGGGACAGAGAAATACCCCTGAATTCGTGCCCGGAAGCCCGAG 827

Qy 158 LeuLeuValPheIleAsnProPheGlyLysGlyGlnGlyLysArgIleTyrGluArg 177
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
828 CTGCTCATATGTTCAATCCCTTTGGGGGGGGGCGCTGGCGCTGGCAGCGCTGTATGAC 887

Qy 178 LysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIlelleValThrGluHis 197
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
888 CACGTGGTGCATATCTCTGAAGCTGGGCTGCTTCAACCTCATACACAGACGAGA 947

Qy 198 AlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleVal 217
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
948 CAGAACCATCGCCGTCGTGAGGGGTTAAGCTGAGTGGAGTGGGAGGCGATTGTC 1007

Qy 218 CysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGln 237
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1008 ACTGTGCTGAGACGGGCTCTTACGAGGTGCTGAATGGCTCCTTGATCGG----- 1061

Qy 238 ArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArg 257
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



```
Db 361 AACATTATGCTGCTATGAGCAGGTCAACAAATGAAGACCTCTCTGACCAACTGCACGCTA 420
Qy 286 IleValValGlyAspSerLeuAla--MetAspValSerSerValHisHisAsnSerThr 304
Db 421 TTGCTGTGCCCGCGCTGCTACCACTGAACCTGTCTCTGACACAGGCTTCGGGG 480
Qy 305 LeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleLeuLysAsp 324
Db 481 CTGGCCCTCTCTCTGTGCTCAGCCTGGCCCTGGGGCTTCATCTGTGTGATGTGGACCTAGAG 540
Qy 325 SerGluLysLysArgTyrPLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPhe 344
Db 541 AGTGAGAAGTATCGCGCTCTGGCGAGATGCGCTTCACTCTGGGCACCTCTCTCGCTG 600
Qy 345 LeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGly 364
Db 601 GCACCCCTGGCCACTACCGCGCGCGACTGGCCCTACCTCCCTGTAGGAAGA---GTGGGT 657
Qy 365 SerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGln 384
Db 658 TCCAAGACACTGCTCCCC-----GTTGTGTGTCACGAGGCGCCGGTA 702
Qy 385 GlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal--- 403
Db 703 GATGCACAC-----CTTGTGCCACTGGAG-----GAGCCAGTGCCT 738
Qy 404 GluGluTrpGlnValValCysGlyLys-----PheLeuAlaIleAsnAlaThr 419
Db 739 TCTCACTGGACAGTGGTGGCCGACGAGACTTTGTGTAGTCTGTCGGCACTGCTGCACCTG 798
Qy 420 AsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGly 439
Db 799 CACTGGCAGTGAATCTTTGCTGCACCCATGGCCGCTGTGCAGCT----- 846
Qy 440 AspGlySerSerAspLeuIleLeuLysLys---CysSerArgPheAsnPheLeuArg 458
Db 847 ---GGCGTCATGCATCTGTCTACGTGCGGCGGAGTGTCTCTGTCATGCTGCTGCGC 903
Qy 459 ---PheLeuIleArgHisThrAsnGlnAspGlnPheAspPheThrPheValGluVal 477
Db 904 CTCCTCTTGGCCATGGAGAGGAGGATGAGTATGAGTATGAGTATGAGTATGATAT 963
Qy 478 TyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeu 497
Db 964 GTGCCCGTGGTGGCTTCGCG-----TTGGAGCCCAAGGAT----- 999
Qy 498 LysGluGlyLysLysLysArgPheGlyHisIleCysSerSerHisProSerCysCys 517
Db 1000 -----GGGAAGGTGTGTTTGA----- 1017
Qy 518 ThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGlu 537
Db 1018 -----GTGATGGGAATTGATGGTTAGCGAGGCGCGTCGAG 1053
Qy 538 ValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnPro--- 556
Db 1054 GGCCAGGTGACCCAACTACTTCTGATGTGTGAGCGGTGCGTGGAGCCCCCGCCGAGC 1113
Qy 557 ---LysPro 558
Db 1114 TGAAGGCC 1122
RESULT 11
US-09-205-258-90
; Sequence 90, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007F1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
```

```
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
```

```

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 90
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (12)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (123)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1522)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1527)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-205-258-90

Alignment Scores:
Pred. No.: 2,258-15 Length: 1533
Score: 257.50 Matches: 113
Percent Similarity: 41.75% Conservative: 64
Best Local Similarity: 26.65% Mismatches: 167
Query Match: 8.51% Indels: 80
DB: 4 Gaps: 18

US-10-631-958-11 (1-562) x US-09-205-258-90 (1-1533)

Qy 154 ArgProLysHisLeuLeuValPheIleAenProPheGlyGlyLeuGlnGlyLysArg 173
Db 109 CGGCCCTCCGGTGNCTGGTCTGCTGAACCCGCGCGGCAAGGCGCAAGCCTTGCG 168
Qy 174 IleTyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIle 193
Db 169 CTCITCCGGAGTCACTGCAGCCCTTTTGGCTGAGGCTGAATCTCTTCCACGCTGATG 228
Qy 194 ValThrGluHisAlaAenGlnAlaLysGluThrLeuTyrGluIleAenIleAspLysTyr 213
Db 229 CTCACGTAGCGCGGCAACACGCGCGGAGTCTGGTGGTGGAGGAGCTGGCGCGCTGG 288
Qy 214 AspGlyIleValCysValGlyAspGlyMetPheSerGluValLeuHisGlyLeuIle 233
Db 289 RACGCTCTGGTGTGATGTYTGAGACGGGCTGATGCAGAGGCTGGTGAACGGGCTTCAT 348
Qy 234 GlyArgThrGlnArgSerAlaGlyValAaspGlnAenHisProArgAlaValLeuValPro 253
Db 349 GGA-----GCGGCTGACTGGGAGACCCCATCCAGAGCC-----CCT 387
Qy 254 -SerSerLeuArgIleGlyIleIleProAlaGlySerThrAspCysValCys----- 270
Db 388 GTGTAGCCTC-----CCAGCAGGCTCTGGCAACGCGCTGGCAGCTTCCTT 432
Qy 271 -----TyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHi 285
Db 433 RAACATTATGCTGGCTATGACGAGGTCAACCAATGAAGACCTCTCGACCACTGCAGCCT 492
Qy 285 sIleValValGlyAspSerLeuAla---MetAspValSerValHisHisAenSerTh 304
Db 493 ATTGCTGTGGCGCGGCTGTGTACCATGAACTGTCTGTCTGACACGCGCTTGGG 552

```

```

Qy 304 rLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAs 324
Db 553 GCTCGCGCTCTTCTCTGTCTGCTCAGCCCTGGCTTCAATGCTGATGGACCTAGA 612
Qy 324 pSerGluLysLysArgTrpLeuGlyLeuAlaAargTyrAspPheSerGlyLeuLysThrPh 344
Db 613 GAGTGAGAGTATCGGGCTCTGGGGAGATGCGGTCTCACTCTGGGCACCTTCTCTGGCTCT 672
Qy 344 eLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGl 364
Db 673 GGCAGCCCTCGCACCTACCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 729
Qy 364 ySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGl 384
Db 730 TTCAAGACACCTGCCTCCCTCCCT-----GTTGTGTCTCCAGCAGGCGCCCGGT 774
Qy 384 nGlnLeuGluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaAspVal-- 403
Db 775 AGATGCACAC-----CTTGTGCCACTGGAG-----GAGCCAGTGCC 810
Qy 404 -GluGluTrpGlnValValCysGlyLys-----PheLeuAlaIleAenAlaTh 419
Db 811 CTCTCACTGGACAGTGGTGGCGGACGAGACTTTGTGCTAGTCTGCTGCTGCTGCTGCTGCT 870
Qy 419 rAenMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGl 439
Db 871 GCACCTGGGCGAGTGAGATGTTTGTGTCACCCCATGGCGCGCTGTCAGCT----- 919
Qy 439 yAspGlySerSerAspLeuIleLeuIleArgLys---CysSerArgPheAenPheLeuAr 458
Db 920 ----GGCGTCATGCATCTGTTCTAGTGGCGGGGAGTGTCTGTCGCATGCTGCTGCG 975
Qy 458 g---PheLeuIleArgHisThrAenGlnAenGlnAenGlnPheAspPheThrPheValGluVa 477
Db 976 CCTCTTCTGGCCATCGAGAAGGCGCAGCATATGAGTATGAATGCCCTACTTGGTATA 1035
Qy 477 lTyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLe 497
Db 1036 TGTGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1072
Qy 497 uLysGluGlyGlyLysLysArgPheGlyHisIleCysSerSerHisProSerCysCy 517
Db 1073 -----GGGAAAGGTGTGTTTGA----- 1090
Qy 517 sThrValSerAenSerSerTrpAenCysAspGlyGluValLeuHisSerProAlaIleGl 537
Db 1091 -----GTGGATGGGAATTGATGTTAGCGAGGCGCGTGCA 1125
Qy 537 uValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAenPro-- 556
Db 1126 GGGCAGGTGCACCAAACTACTTCTGGATGGTCTGAGCGGTGCTGAGCGCCCGCCAG 1185
Qy 557 ----LysPro 558
Db 1186 CTGGAAGCCC 1195

RESULT 12
US-09-970-516-5
; Sequence 5, Application US/09970516
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; TITLE OF INVENTION: polynucleotides encoding sphingosine kinases
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1149
; TYPE: DNA
; ORGANISM: Mus musculus

```

[illegible]





```
Db 666 TCTGCAACGGCTGGCAGCTTCCTTGAACCAATTATGCTGGCTATGAGCAGGTCAACAAAT 725
Qy 277 SerAepAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAla--MetAsp 295
Db 726 GAAGACCTCTCGAACCACTGACAGCTATGCTGTGCGCCGCGCTGTCTACCCATGAAC 785
Qy 296 ValSerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyr 315
Db 786 CTGCTGCTCTGCACACGGCTCGGGGCTGGCCCTCTCTCTGTGCTCAGCCTGGCCCTGG 845
Qy 316 GlyPheTyrGlyAspIleIleLysAspSerGlnLysLysArgTrpLeuGlyLeuAlaArg 335
Db 846 GGCCTTCATGCTGACACGGCTCGGGGCTGGCCCTCTCTCTGTGCTCAGCCTGGCCCTGG 905
Qy 336 TyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSer 355
Db 906 TTCACTCTGGGCACCTTCCCTCGCTGTCGACGCCCTGGCCACTACCGCGCGGACTGGCC 965
Qy 356 PheLeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGly 375
Db 966 TACTCTCTCTAGGAGA--GTGGGTCCAGACACCTGCTCCGCC----- 1010
Qy 376 CysPheValCysArgGlnSerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyr 395
Db 1011 --GTTGTGTCCAGCAGCGCCGCTAGATGCACAC-----CTTGTG 1049
Qy 396 GlyLeuGluAlaAlaGluAspVal---GluGluTrpGlnValValCysGlyLys----- 412
Db 1050 CCACTGGAG-----GAGCCAGTCCCTCTCACTGGACAGTGGTGGCCGACGAGGACTTT 1103
Qy 413 -----PheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArg 430
Db 1104 GTGCTAGTCTCTGGCACTGCTGCATCGCACCTGGCAGTGCAGTGCAGTTCGTGACCCATG 1163
Qy 431 GlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuLeuLeuLeuArgLys 450
Db 1164 GCGCGCTGTGCAGCT-----GCGCGTCATGCATCTGTTCTACGTGCGGCGG 1208
Qy 451 ---CysSerArgPheAsnPheLeuArg---PheLeuIleArgHisThrAsnGlnGlnAsp 468
Db 1209 GGAGTGTCTCTGCGCATGCTGCTGCGCTCTTCCTGGCCATGGAGAGGGCAGCATATG 1268
Qy 469 GlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPheThrSerLys 488
Db 1269 GAGTATGAATGCCCTACTTGTATATGTGCGCGTGGTGGCTTCGCG----- 1316
Qy 489 HisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPheGlyHisIle 508
Db 1317 --TTGGAGCCCAAGGAT-----GGGAAAGGTGTGTTTGCA----- 1349
Qy 509 CysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsnCysAspGly 528
Db 1350 -----GTGGATGGG 1358
Qy 529 GluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuValArgLeuPhe 548
Db 1359 GAAATGATGTTAGCGAGCGGTGCGAGCGGCGGAGGTGCACCCAACTACTTCTGCGATGTC 1418
Qy 549 AlaArgGlyIleGluGluAsnPro-----LysPro 558
Db 1419 AGCGGTTGCTGGAGCCCGCCGCGCTGGAAGGCC 1454
```

## RESULT 15

```
US-09-614-221A-399
; Sequence 399, Application US/09614221A
; Patent No. 6723837
; GENERAL INFORMATION:
; APPLICANT: Karunanandaa, Balasulojini
; APPLICANT: Yu, Jaehyuk
; APPLICANT: Kishore, Ganesh M.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED
; WITH STEROL SYNTHESIS AND METABOLISM
```

```
; FILE REFERENCE: 16516.075
; CURRENT APPLICATION NUMBER: US/09/614,221A
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/142,981
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 626
; SEQ ID NO 399
; LENGTH: 1875
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-614-221A-399

Alignment Scores:
Pred. No.: 1.37e-10 Length: 1875
Score: 210.50 Matches: 66
Percent Similarity: 44.62% Conservative: 46
Best Local Similarity: 26.29% Mismatches: 102
Query Match: 6.96% Indels: 37
DB: 4 Gaps: 6
```

US-10-631-958-11 (1-562) x US-09-614-221A-399 (1-1875)

```
Qy 143 ThrLeuArgGluMetLeuGluLysLeu-----ThrSerArgProLysHisLeuLeu 159
Db 631 ACTGTGGAAGAAATCTCTGGAAGAAATTTATGAAAAATTCGAAGAGAAACAGATCGATATTA 690
Qy 160 ValPheIleAsnProPheGlyGlyGlyGlnGlyLysArgIleTyrGluArgLysVal 179
Db 691 GTCAATTATATCCCAACGGTGTAAAGGTACTGCTAAAAATTTATTTCTCGACAAAAGCA 750
Qy 180 AlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsn 199
Db 751 AGGCCAATACTAGTGGAAAGTGGCTGCAAAATAGAAATTCATACACAAAATATGCCCGT 810
Qy 200 GlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysVal 219
Db 811 CACGCCATCGATATTCGCAAGATTTAGATATCAGCAAAATACGATACCATTTGCGATGGCC 870
Qy 220 GlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSer 239
Db 871 TCGGTGATGTTATTCATACGAAAGTAATTAATGGGCTTTATAGAAAGCCCGACAGA--- 927
Qy 240 AlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGly 259
Db 928 -----GTGGATGCGTTCATATAAACTAGCCGTAACCTCAG----- 960
Qy 260 IleIleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAla 279
Db 961 ---CTACCTTTCGCGTTTCAGGAAATGCTATGAGCATTTTCATGTTCATCGACAAATAACCCA 1017
Qy 280 GluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspVal----- 296
Db 1018 TCGTACCGCGCTCTGTCCTTGTCAATCCATTCGAAACAAGATAGACTTAATGTGTGT 1077
Qy 297 SerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuGlyTyrGly 316
Db 1078 TCCCAACCTCTTATATGAACGAATGCGCAAGATTATCTCTTTTGTAGTCAGACAGCTACGGC 1137
Qy 317 PheTyrGlyAspIleIleLysAspSerGlnLysLysArgTrpLeuGlyLeuAlaArgTyr 336
Db 1138 GTTATTCAGAAATCTGATATTAACACATTCATCAGATGGATGGGTCCCGTTAGGTGT 1197
Qy 337 AspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPhe 356
Db 1198 AATTGGGT-----GTAGCATTCACATTATCCAAAGT----- 1230
Qy 357 LeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCys 376
Db 1231 -----AAGAAATATCCCTCGCAAGTTTTCGTC 1257
Qy 377 PheValCysArgGlnSerLysGlnLeuGlu 387
Db 1258 AAATATGCTGCCAAATCTAAAAAGGAATAAAA 1290
```

Search completed: September 6, 2005, 02:35:04  
Job time : 367.875 secs



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2005, 23:41:54 ; Search time 1593.07 Seconds  
(without alignments)  
2324.913 Million cell updates/sec

Title: US-10-631-958-11  
Perfect score: 3025  
Sequence: 1 HEANGPAPLGVRAAPPWRT.....QLVRLFARGIENKPKDPSHS 562

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 7338684 seqs, 3274456166 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US10631958/runat\_02092005\_165813\_4470/app\_query.fasta\_1.1941  
-DB=Published Applications NA -OPMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62  
-TRANS=human40.cdi -LISTS=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10631958 @CGN 1 1 1588 @runat\_02092005\_165813\_4470  
-NCPV=6 -ICPU=3 -NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- Published Applications NA:\*
- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
  - 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
  - 3: /cgn2\_6/ptodata/1/pubpna/US05\_NEW\_PUB.seq:\*
  - 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
  - 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
  - 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
  - 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
  - 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
  - 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
  - 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
  - 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
  - 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
  - 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
  - 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
  - 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
  - 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
  - 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
  - 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*
  - 19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq:\*
  - 20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq:\*
  - 21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq:\*
  - 22: /cgn2\_6/ptodata/1/pubpna/US10J\_NEW\_PUB.seq:\*
  - 23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq:\*
  - 24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
  - 25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
  - 26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3025	100.0	1740	18	US-10-262-511-39	Sequence 39, Appl
2	3025	100.0	4413	10	US-09-969-896-16	Sequence 16, Appl
3	3025	100.0	4413	19	US-10-631-958-16	Sequence 16, Appl
4	3012	99.6	4429	19	US-10-618-941-55	Sequence 55, Appl
5	3004	99.3	4463	16	US-10-315-597A-1	Sequence 1, Appli
6	2931.5	96.9	4432	17	US-10-120-988-148	Sequence 148, App
7	2888	95.5	1614	10	US-09-969-896-9	Sequence 9, Appli
8	2888	95.5	1614	19	US-10-631-958-9	Sequence 9, Appli
9	2649.5	87.6	1840	9	US-09-784-810A-5	Sequence 5, Appli
10	2649.5	87.6	1840	22	US-10-876-281-5	Sequence 1, Appli
11	1640.5	54.2	979	10	US-09-969-896-1	Sequence 1, Appli
12	1640.5	54.2	979	19	US-10-631-958-1	Sequence 247, App
13	1156	38.2	817	19	US-10-115-635-247	Sequence 8, Appli
14	1008	33.3	550	19	US-09-969-896-8	Sequence 8, Appli
15	1008	33.3	550	19	US-10-631-958-8	Sequence 4, Appli
16	811	26.8	474	10	US-09-969-896-4	Sequence 4, Appli
17	811	26.8	474	19	US-10-631-958-4	Sequence 7, Appli
c 18	656	21.7	522	9	US-09-784-810A-7	Sequence 7, Appli
c 19	656	21.7	522	22	US-10-876-281-7	Sequence 176413,
20	616.5	20.4	2657	20	US-10-425-115-176413	Sequence 34530, A
21	573.5	19.0	2079	18	US-10-425-114-34530	Sequence 86211, A
22	573.5	19.0	2079	20	US-10-425-115-86211	Sequence 5, Appli
23	525	17.4	329	10	US-09-969-896-5	Sequence 5, Appli
24	525	17.4	329	19	US-10-631-958-5	Sequence 92801, A
c 25	480.5	15.9	2189	19	US-10-437-963-92801	Sequence 46, Appl
26	437.5	14.5	564	22	US-10-477-445-46	Sequence 30650, A
27	410	13.6	1833	19	US-10-437-963-30650	Sequence 64890, A
28	392.5	13.0	2165	20	US-10-425-115-64890	Sequence 47396, A
29	389	12.9	2084	18	US-10-424-599-47396	Sequence 34806, A
30	388	12.8	1869	18	US-10-425-114-34806	Sequence 5271, Ap
31	371	12.3	339	9	US-09-783-590-5271	Sequence 23002, A
32	363	12.0	3362	20	US-10-425-115-29002	Sequence 59419, A
33	347	11.5	2014	18	US-10-424-599-59419	Sequence 64892, A
34	344.5	11.4	1512	19	US-10-425-115-64892	Sequence 58701, A
35	344.5	11.4	1916	19	US-10-437-963-58701	Sequence 9, Appli
36	343.5	11.4	382	9	US-09-784-810A-9	Sequence 34081, A
37	343.5	11.4	382	22	US-10-876-281-9	Sequence 95148, A
38	339	11.2	1774	18	US-10-425-114-34081	Sequence 77896, A
39	331.5	11.0	1386	19	US-10-437-963-95148	Sequence 25, Appl
40	330.5	10.9	3008	17	US-10-348-052-25	Sequence 25, Appl
41	329.5	10.9	2609	19	US-10-622-011-25	Sequence 32139, A
42	329.5	10.9	2609	19	US-10-425-114-32139	Sequence 97849, A
43	326.5	10.8	1857	18	US-10-425-115-97849	Sequence 79424, A
44	326.5	10.8	2247	20	US-10-425-115-97849	
45	316.5	10.5	2284	20	US-10-425-115-79424	

ALIGNMENTS

RESULT 1  
US-10-262-511-39  
; Sequence 39, Application US/10262511  
; Publication No. US20040038223A1  
; GENERAL INFORMATION:  
; APPLICANT: Smithson, Glenna  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen

```
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Raestelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; PRIOR FILING DATE: 2003-05-28
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuroSeqList version 0.1
; SEQ ID NO 39
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (76)..(1686)
US-10-262-511-39

Alignment Scores:
Pred. No.: 0 Length: 1740
Score: 3025.00 Matches: 562
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-631-958-11 (1-562) x US-10-262-511-39 (1-1740)

QY 1 HisGluAlaAAsenGlyProAlaProLeuGlyValArgAlaProProAlaTrpArgThr 20
DB 1 CACGAGGCGGTAAACGGTCCGGCGCCCTCGCGCGTCCGCCGCCGCCAGCCCTGGCGGACG 60
QY 21 SerProAlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrp 40
DB 61 AGCCCGCGGGAGATGGGGCGGACGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
QY 41 VallysGlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrp 60
121 GTGAAGCAGCAGCGCTCGCGCGTGGAGCCCGGGGCTCTGCTGCGCTGGTGG 180
61 ArgSerProGlyProGlyAlaGlyAlaProGlyAlaAlaPalaaCysSerValProValSer 80
181 CGGAGCCCGGGGCGCGGAGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
81 GluIleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrp 100
241 GAGATCATCGCGTTGAGGAAACAGAGCTTCAACGGGAAACATCAAGGAGTGGAAATGG 300
101 GlnLysMetGluLysProTrpAlaPheThrValHisCysValLysArgAlaArgHis 120
301 CAGAAATGGAAGACCTTACGCTTTTACAGTTTCACTGTGTAAAGAGACGACGCGCAC 360
121 ArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGlnLeuCysHisLeuTrp 140
361 CGCTGGAGTGGCGCAGGTGACTTTCTGGTGTCCAGAGGAGCAGTGTGTCACTTGTGG 420
141 LeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuVal 160
421 CTGCAGACCTTGGCGGAGATGCTGGAGAGCTGACGTCAGACCCAAAGCATTTTACTGGTA 480
161 PheIleAsnProPheGlyLysGlyGlnGlyLysArgIleTrpGluArgLysValAla 180
481 TTTATCAACCGTTTGGAGAAAGGACAAGGCAAGCGGATATATGAAGAAAAAGTGCA 540
181 ProLeuPheThrLeuAlaSerIleThrThrAspIleValThrGluHisAlaAsnGln 200
541 CCACTGTTTCACTTGGCTTCATCCACTGACATCATCTGTTACTGAACATGCTTAATCAG 600
201 AlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGly 220
601 GCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGACGCGCATCGTCTGTGCGG 660
221 GlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAla 240
661 GGAGATGGTATGTTTTCAGCGAGTCTGACGGTCTGATTGGAGGACGACGAGAGCGCC 720
241 GlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIle 260
721 GGGGTCCAGCAGAACACCCCGGGCTGTGTGGTCCCGAGTAGCCTCCGAGTTGGAAATC 780
261 IleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGlu 280
781 ATTCGCCGAGGTCACAGGACTGCGTGTGTACTCCACCGTGGGACACGACGACGAGAA 840
281 ThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHis 300
841 ACCTCGCGCTGCATATCGTTGTTGGGAGCTCGCTGGGCCATGGATGTGTCTCAGTCCAC 900
301 HisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAsp 320
901 CACAACAGCACACTCTTCGCTACTCCGTGTCCCTGCTGGGCTACCGGCTTCTACGGGAC 960
321 IleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGly 340
961 ATCATCAGGACAGTGGAGAGAACCGTGGTGGGCTTCCAGATACGACTTTTTCAGGT 1020
341 LeuLysThrPheLeuSerHisHisCysTyrGlyThrValSerPheLeuProAlaGln 360
1021 TTAAGAACCTTCTCTCCACCACTGCTATGAAGGAGCAGTGTCTCTCTCTCTCTCTCTCT 1080
361 HisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArg 380
1081 CACCGTGGGATCTCCAAGGGATAGAGAGCCCTGCGGGGAGGATGCTTTGTTTGAGG 1140
381 GlnSerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAla 400
1141 CAAAGCAAGCAGCAGCTGGAGGAGGAGCAGACAGAAAGACACTGTATGTTTGGAAAGTCG 1200
401 GluAspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsn 420
```

```
Db 1201 GAGGACGTGGAGGAGTGGCAAGTCTGTGTGGAGATTCTTGGCCATCAATGCCACAAAC 1260
Qy 421 MetSerCysAlaCysArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAsp 440
Db 1261 ATGCTCTGTGTTGTGCGCGAGCCCAAGGGGCTCTCCCGGCTGCCCACTTGGAGAC 1320
Qy 441 GlySerSerAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 460
Db 1321 GGGTCTTCTGACCTCATCTCATCCGAAATGCTCCAGGTCAATTTCTGAGATTTC 1380
Qy 461 IleArgHisThrAsnGlnAspGlnPheAspPheThrPheValGluValValArgVal 480
Db 1381 ATCAGGCACACCAACAGCAGGACCACTTGTGACCTTCACTTTTGTGAGTTTATCGCGTC 1440
Qy 481 LysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGly 500
Db 1441 AAGAAATCCAGTTTACGTCGAAGCACATGGAGGATGAGGACAGCCCTCAAGGAGGGG 1500
Qy 501 GlyLysLysArgPheGlyHisLysCysSerSerHisProSerCysCysThrValSer 520
Db 1501 GGGAAAGAGCGCTTTGGGACATTTGCGAGCAGCCACCCCTCTCTGCTGTCACCGTCTCC 1560
Qy 521 AsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgVal 540
Db 1561 AACAGCTCTGGACTGCGAGCGGGAGGTCTGCGACAGCCCTGCCATCGAGGTCAAGTCA 1620
Qy 541 HisCysGlnLeuValArgLeuPheAlaArgGlyLysGluGluAsnProLysProAspSer 560
Db 1621 CACTGCCAGCTGGTTGCGACTCTTTGCAAGGAAATGAAGAAATCCGAAAGCCAGACTCA 1680
Qy 561 HisSer 562
Db 1681 CACAGC 1686
```

## RESULT 2

```
; Sequence 16, Application US/09969896
; Publication No. US20030125533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 4413
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-16
```

```
Alignment Scores:
Pred. No.: 0 Length: 4413
Score: 3025.00 Matches: 562
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
```

US-10-631-958-11 (1-562) x US-09-969-896-16 (1-4413)

```
Qy 1 HisGluAlaAlaAsnGlyProAlaProLeuGlyValArgAlaProProAlaTrpArgThr 20
Db 1 CACGAGGCGCGTTAACGGTCCGCGCGCCCTCGCGCTCGCGCGCCGCCCGCCAGCTGGCGGACG 60
Qy 21 SerProAlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrp 40
```

```
Db 61 AGCCCGCGCGCGAGATGGGGGCGACCGGGGCGCGGAGCCGCTGCAATCCGTGCTGTGG 120
Qy 41 ValLysGlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrp 60
Db 121 GTGAAGCAGCAGCCTCGCGCTGAGCGCCGCGCGGCTCTGCTGCGCTGTGTGG 180
Qy 61 ArgSerProGlyProGlyAlaGlyValProGlyAlaAlaPalaCysSerValProValSer 80
Db 181 CGGAGCCCGGGGCGCGGAGCGCGCGCGCGCGGATGCGCTCTCTGTGCTGCTGTATCT 240
Qy 81 GluIleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrp 100
Db 241 GAGATCATCGCGCTTGAGGAAACAGAGCTTCACGGGAAACATCAAGGCAGTGGAAATGG 300
Qy 101 GlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgHis 120
Db 301 CAGAAATGGAAAGCCCTTACGCTTTTACAGTTTCACTGTGTAAAGAGAGACGACGCGCAC 360
Qy 121 ArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrp 140
Db 361 CGCTGGAGAGTGGGCGCAGGTGACTTCTGTGTGTCCAGAGGAGCAGCTGTGTCTACTTGG 420
Qy 141 LeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuVal 160
Db 421 CTGCAGACCCCTGCGGAGATGCTGGAGAGCTGACGCTCCAGACCAAGCATTTACTGGTA 480
Qy 161 PheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAla 180
Db 481 TTTTATCAACCGTTTGGAGGAAAGGACAAAGGCAAGCGGATATATGAAGAAAGTGGCA 540
Qy 181 ProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGln 200
Db 541 CCACCTGTTCACTTTAGCTTCCATCACCACCTGACATCATCGTTACTGAACATGCTAATCAG 600
Qy 201 AlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGly 220
Db 601 GCCAAGGAGACTCTGTATGAGATTAAATACAGCAAAATACGACCGGATCGTCTGTGTCGGC 660
Qy 221 GlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAla 240
Db 661 GGAGATGTTATGTTTCAGCAGGTGCTGCACGGTCTGATTGGGAGGACGCGAGAGCGCC 720
Qy 241 GlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIle 260
Db 721 GGGGTGACACCAAGAACCCCGCGGCTGTGCTGCTCCCGAGTAGCCTCCGATTGGAAATC 780
Qy 261 IleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGlu 280
Db 781 ATTCGCGAGGGTCAACGGACTGCTGTGTACTTCCACCGTGGGACGACGAGCGCAGAA 840
Qy 281 ThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHis 300
Db 841 ACCTCGCGCTGCATATCGTTGTTGGGACTCGCTGGGCAATGGATGTGCTCCTCAGTCCAC 900
Qy 301 HisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAsp 320
Db 901 CACAACAGCACACTCTCTCGCTACTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Qy 321 IleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGly 340
Db 961 ATCATCAGGACAGTGGAGAAACGGTGGTGGGTCTTGGCAGATACGACTTTTTCAGGT 1020
Qy 341 LeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGln 360
Db 1021 TTTAAAGACCTTCTCTCCACCACTGCTATGAAGGGGACAGTGTCTTCTCTCCCTGACAA 1080
Qy 361 HisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArg 380
Db 1081 CACACGCTGGGATCTCCAAGGGATAGGAAGCCCTGCGCGGCGAGGATGCTTTGTTTGCAGG 1140
Qy 381 GlnSerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAla 400
Db 1141 CAAAGCAAGCAGCAGCTGGAGGAGGAGCAGAAAGACACTGTATGTTTGGAAAGCTCGG 1200
```





Db 1081 CACACGGTGGGATCTCAAGGAGATAGGAAGCCCTCGCGGCAGGATGCTTTGTTGTCAGG 1140  
Qy 381 GlnSerLysGlnGlnLeuGluGlnGlnLysAlaLeuTyrGlyLeuGluAlaAla 400  
Db 1141 CAAAGCAAGCAGCAGCTGGAGAGAGAGCAAGAAAGCACTGTATGGTTGGAAAGCTGG 1200  
Qy 401 GluAspValGluGluTyrGlnValValCysGlyLysPheLeuAlaAlaAlaThrAsn 420  
Db 1201 GAGGACGTGGAGAGTGGCAAGTCTGTGTGGGAAGTTTCTGCCCATCATGCCCACAAAC 1260  
Qy 421 MetSerCysAlaCysArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAsp 440  
Db 1261 ATGTCTGT 1320  
Qy 441 GlySerSerAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 460  
Db 1321 GGGTCTTCTGACCTCATCTCATCGGAAATGCTCCAGGTTTCAATTTTCTGAGATTTCTC 1380  
Qy 461 IleArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgVal 480  
Db 1381 ATCAGGCACCAACAGCAGGACCAAGTTTGACCTTTTGTGGAAGTTTATCGCGTC 1440  
Qy 481 LysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGly 500  
Db 1441 AAGAAATTCAGTTTACGTTCGAAGCACATGGAGGATGAGGACAGCCTCAAGGAGGG 1500  
Qy 501 GlyLysLysArgPheClyHisIleCysSerSerHisProSerCysCysThrValSer 520  
Db 1501 CGGAAGAAGCGCTTTGGGCACATTTGGCAGCAGCACCCTCTCTGTGTGTGTGTGTGTGT 1560  
Qy 521 AsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaAlaGluValArgVal 540  
Db 1561 AACAGCTCTGGAACCTGCGCGGAGGTCCTGCACAGCCCTGCGCCATCGAGGTCAAGTTC 1620  
Qy 541 HisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSer 560  
Db 1621 CACTGCCAGCTGTTTCGACTCTTTGCACGAGGAATGAAGAGAAATCGNAGCCACTCA 1680  
Qy 561 HisSer 562  
Db 1681 CACAGC 1686

## RESULT 4

US-10-618-941-55  
; Sequence 55, Application US/10618941  
; Publication No. US2004019792A1  
; GENERAL INFORMATION:  
; APPLICANT: WHYTE, DAVID  
; APPLICANT: MANNING, GERARD  
; APPLICANT: CAENEPEEL, SEAN  
; TITLE OF INVENTION: NOVEL KINASES  
; FILE REFERENCE: 034536-0321  
; CURRENT APPLICATION NUMBER: US/10/618,941  
; CURRENT FILING DATE: 2003-07-15  
; PRIOR APPLICATION NUMBER: 60/395,632  
; PRIOR FILING DATE: 2002-07-15  
; NUMBER OF SEQ ID NOS: 143  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 55  
; LENGTH: 4429  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-618-941-55

## Alignment Scores:

Pred. No.:	0	Length:	4429
Score:	3012.00	Matches:	560
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	99.57%	Indels:	0
DB:	19	Gaps:	0

US-10-631-958-11 (1-562) x US-10-618-941-55 (1-4429)

Qy 3 AlaAlaAsnGlyProAlaProLeuGlyValArgAlaProProAlaTyrArgThrSerPro 22  
Db 23 GCCGCTAACCGTCCGGCGCCCTCGCGCTCGCGCGCCCGCCAGCGTGGCGAGCGAGCCG 82  
Qy 23 AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTyrValLys 42  
Db 83 CGCGCGAGATGGGGCGAGCGGGGGCGGAGCGCGCTGCAATCCGTCGCTGGGGTGAAG 142  
Qy 43 GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTyrTrpArgSer 62  
Db 143 CAGCAGCGCTCGCGCGTGGAGCTCGAGCCCGCGCGGCTCTGCTCGCTGGTGGCGAGC 202  
Qy 63 ProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIle 82  
Db 203 CGCGCGCGCGAGCGCGCGCCCGCGCGGATGCTGCTCTGTGCTCTGTATCTAGATC 262  
Qy 83 IleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLys 102  
Db 263 ATCGCGCTTGAGGAAACAGAGCTTCA CGGGAACATCAAGGAGTGGAAATGGCAGAAA 322  
Qy 103 MetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrp 122  
Db 323 ATGGAAGAAGCCTTACGCTTTTACAGTTCACTGTGTAAAGAGAGCAGCAGCGCACGCTGG 382  
Qy 123 LysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGln 142  
Db 383 AAGTGGCGCAGGTGACTTCTGTGTGTCCAGGAGGAGCAGTGTGTCACTTGTGGTGCAG 442  
Qy 143 ThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIle 162  
Db 443 ACCCTGCGGAGATGCTGGAGAGAGCTGACGTCCAGACCAAGACATTTTACTGGTATTATC 502  
Qy 163 AsnProPheGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeu 182  
Db 503 AACCGTTTGGAGAAAGAGCAAGGCAAGCGGATATATGAAGAAAAAGTGGCACCACTG 562  
Qy 183 PheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLys 202  
Db 563 TTCACCTTAGCCTCCATCACCACCTGACATCATCGTTACTGAACATGCTTAATCAGGCCAAG 622  
Qy 203 GluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyLysAsp 222  
Db 623 GAGACTCTGTATGAGATTACATAGACAAATACAGCGGCATCGTCTGTGTGGCGGAGAT 682  
Qy 223 GlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyVal 242  
Db 683 GGTATGTTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGAGCGCGGGGTC 742  
Qy 243 AspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIlePro 262  
Db 743 GACCAGAACCAACCCCGCGCTGTGTGTCTCCCAAGTAGCCTCCGATTTGGAATCAATCCC 802  
Qy 263 AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer 282  
Db 803 CGAGGTCACAGGACTGCGTGTGTGTACTCCACCCTGGGCGACGAGCGCAGAAACCTCG 862  
Qy 283 AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsn 302  
Db 863 CGCGTGCATATCGTTTGGGAGCTCGCTGCGCCATGGATGTGCTCTCAGTCCACCCACAAC 922  
Qy 303 SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIle 322  
Db 923 AGCACTCTCTCGCTACTCGCTGTCTGCTGTGGCTACGGCTTCTACGGGAGACATCATC 982  
Qy 323 LysAspSerGluLysValLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys 342  
Db 983 AAGCAGTGTGAGAGAAACCGTGGTGGTCTTCCAGATACAGCTTTTACAGTTTAAAG 1042  
Qy 343 ThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr 362  
Db 1043 ACCTTCTCTCCACCACCTGCTATGAGGGGACAGTGTCTTCTCTCTCTCTCTCTCTCTCT 1102

Qy	363	ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer	382
Db	1103	GTGGGATCTCCAAAGGGATAGAAAGCCCTCGCGGGCAGGATGCTTTGTTGTCAGCGAAAGC	1162
Qy	383	LysGlnGlnLeuGluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp	402
Db	1163	AAGCAGCAGCTGGAGGAGGAGCAGAGAAAGCACTGTATGTTTGGAAAGCTGGGAGGAC	1222
Qy	403	ValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSer	422
Db	1223	GTGGAGGAGTGGCAAGTCGTCGTGGTGGAAAGTTCTTGCGCCATCAATGCCACAAACATGTCC	1282
Qy	423	CysAlaCysArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer	442
Db	1283	TGTGCTTTGTGCGCGAGGCCCGAGGGGCTCTCCC CGGCTGCCCACTTGGGAGACGGGTCT	1342
Qy	443	SerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArg	462
Db	1343	TCTGACCTCATCCTTCATCCGGAAATGCTCAGGTTCAATTTTTCGAGATTTTCTCATCAGG	1402
Qy	463	HisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLys	482
Db	1403	CACACCAACAGCAGGAGGACCAAGTTTGACTTTCCTTTGTAAGTTTATCGGTCACAGAAA	1462
Qy	483	PheGlnPheThrSerLysHisMetGluAspGluaspSerAspLeuLysGluGlyLys	502
Db	1463	TTCCAGTTTATCGTCCGAAGCATCGAGGATGAGACACGACCTCAAGAGGGGGGGAAG	1522
Qy	503	LysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSer	522
Db	1523	AAGCGCTTTTGGGCACATTTTGCAGCAGCCACCCCTCTGCTGCTGCACCGTCTCCACAGC	1582
Qy	523	SerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCys	542
Db	1583	TCCTGGAACTGCCAGCGGGAGGTCTTCGACAGGCCCTCGCCATCGAGGTCAGAGTCCACTGC	1642
Qy	543	GlnLeuValArgLeuPheAlaArgGlyIleGluAsnProLysProAspSerHisSer	562
Db	1643	CAGCTGGTTTGACTTTTTCACAGAGGAATTTGAAGAAATCCGAAGCCGAGACTCACACAGC	1702

## RESULT 5

```

US-10-315-597A-1
; Sequence 1, Application US/10315597A
; Publication No. US20030162206A1
; GENERAL INFORMATION:
; APPLICANT: Suglura, Masako
; APPLICANT: Kono, Keita
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Ceramide Kinase and DNA Encoding It
; FILE REFERENCE: 02658CIP/HG
; CURRENT APPLICATION NUMBER: US/10/315,597A
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: JP 2000-178039
; PRIOR FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4463
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 3371
; OTHER INFORMATION:
US-10-315-597A-1

Alignment Scores:
Pred. No.: 0 Length: 4463
Score: 3004.00 Matches: 558
Percent Similarity: 99.82% Conservative: 1
Best Local Similarity: 99.64% Mismatches: 1
Query Match: 99.31% Indels: 0
DB: 16 Gaps: 0

```

Db 1075 ACCTTCCTCCACACCTGCTATGAGGAGCAGTGTCTTCTCCTCCCTGCACAAACACAGG 1134  
 QY 363 ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer 382  
 Db 1135 GTGGGATCTCCAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTGCAGGCAAGC 1194  
 QY 383 LysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp 402  
 Db 1195 AAGCAGCAGCTGGAGAGAGAGAGAGAAAGCACTGTATGGTTTGAAGCTCCGGAGGAC 1254  
 QY 403 ValGluGluTrpGlnValValCysGlyLysPheLeuAlaAlaAsnAlaThrAsnMetSer 422  
 Db 1255 GTGGAGGAGTGGCAAGTCTGTGTGGAGATTTCTGGCCATCAATGCCACAAACATGTCC 1314  
 QY 423 CysAlaCysArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer 442  
 Db 1315 TGTGCTTGTGCGCGGAGCCAGGGGCTCTCCCGGCTGCCCACTTGGGAGACGGTCT 1374  
 QY 443 SerAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 462  
 Db 1375 TGTGACCTCATCTCATCCGGAATGCTCAAGTTCAATTTTCTGAGATTTCTCATCAGG 1434  
 QY 463 HisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLys 482  
 Db 1435 CACACCAACAGCAGGACCACTTTCACCTTCTTGTGAGTTTATCGCTCAAGAAA 1494  
 QY 483 PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuGlyGlyLys 502  
 Db 1495 TTTCCAGTTTACCTCGAAGCACATGAGGATGAGGACGACCTCAAGGAGGGGGGAAG 1554  
 QY 503 LysArgPheGlyHisLysSerHisProSerCysCysCysThrValSerAsnSer 522  
 Db 1555 AAGCCCTTTGGGCACATTTGACAGCACCACCCCTCTGCTGTCACCGTCTCCAAACAGC 1614  
 QY 523 SerTrpAsnCysAspGlyGluValLeuHisSerProAlaGluValArgValHisCys 542  
 Db 1615 TCTTGGAACTGCGATGGGAGGTCTGACACACCTTCCATCGAGTTCAGTCCACTGC 1674  
 QY 543 GlnLeuValArgLeuPheAlaArgGlyLeuGluGluAsnProLysProAspSerHisSer 562  
 Db 1675 CAGCTGGTTCGACTCTTTGCACGAGGAATTTGAAGAGATCCGAAGCCAGACTCACACAGC 1734

## RESULT 6

US-10-120-988-148  
 ; Sequence 148, Application US/10120988  
 ; Publication No. US20030219745A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tang, Y. Tom  
 ; APPLICANT: Goodrich, Ryle  
 ; APPLICANT: Liu, Chenghua  
 ; APPLICANT: Ren, Feiyun  
 ; APPLICANT: Wang, Dunrui  
 ; APPLICANT: Drmanac, Radoje T.  
 ; TITLE OF INVENTION: No. US20030219745A1el Nucleic Acids and  
 ; FILE OF INVENTION: Polypeptides  
 ; FILE REFERENCE: 802CON  
 ; CURRENT APPLICATION NUMBER: US/10/120,988  
 ; CURRENT FILING DATE: 2002-04-11  
 ; PRIOR APPLICATION NUMBER: 09/774,528  
 ; PRIOR FILING DATE: 2001-01-30  
 ; NUMBER OF SEQ ID NOS: 441  
 ; SOFTWARE: pt\_FL\_genes Version 2.0  
 ; SEQ ID NO 148  
 ; LENGTH: 4432  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1129) .. (2817)  
 US-10-120-988-148  
 Alignment Scores: 4.03e-313 Length: 4432  
 Pred. No.:

Score: 2931.50 Matches: 548  
 Percent Similarity: 98.21% Conservative: 2  
 Best Local Similarity: 97.86% Mismatches: 7  
 Query Match: 96.91% Indels: 3  
 DB: 17 Gaps: 1  
 US-10-631-958-11 (1-562) x US-10-120-988-148 (1-4432)  
 QY 6 GlyProAlaProLeuGly-----ValArgAlaProProAlaTrpArgThrSerPro 22  
 Db 1135 GGTTCGCCCTATAGACACAGACAGTCAACAGGAGCTGCAGGCTTGGCGGACGAGCCCG 1194  
 QY 23 AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLys 42  
 Db 1195 CGCGCGAGATGGGGCGGACCGGGCGCGGAGCCGCTGCAATCCGTGCTGTGGGTGAAG 1254  
 QY 43 GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSer 62  
 Db 1255 CAGCAGCGCTGCGCGCTGAGCTGGAGCCCGCGGGCTCTGCTGCGTGTGGTGGGAGC 1314  
 QY 63 ProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluLeu 82  
 Db 1315 CGGGGCGCGAGCGCGGCCCGCGGATGCTCTGCTGCTGTATCTGAGATC 1374  
 QY 83 IleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLys 102  
 Db 1375 ATCGCCGTTGAGGAAACAGACGTTTCACGGGAAACATCAAGGCAGTGGGAAATGGCAGAAA 1434  
 QY 103 MetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrp 122  
 Db 1435 ATGGNAAAGCCTTACGCTTTTACGTTTCTGTGTAAAGAGACGACGACGCGCTGG 1494  
 QY 123 LysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGln 142  
 Db 1495 AAGTGGCGCAGGTGACTTTCTGGTGTCCAGAGGAGCAGCTGTGTCTGTGGTGCAG 1554  
 QY 143 ThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheLeu 162  
 Db 1555 ACCCTGCGGGAGATGCTGGAGAAGCTGACGTCCAGACCAAGCAATTTACTGGTATTATC 1614  
 QY 163 AsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeu 182  
 Db 1615 AACCCGTTTGGAGNAAAGGACAGGCAAGCGGATATATGAAGAAAGTGGCACACTG 1674  
 QY 183 PheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLys 202  
 Db 1675 TTCACCTTAGCCTCCATCACCACCTGACATCATCGTTACTGAACATGCTAATCAGGCCAAG 1734  
 QY 203 GluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAsp 222  
 Db 1735 GAGACTCTGTATGAGATTAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGAT 1794  
 QY 223 GlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyVal 242  
 Db 1795 GGTATGTTACGAGGTGCTGACGCTGATTTGGAGAGACGACGAGAGCGCCGGGTGTC 1854  
 QY 243 AspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIlePro 262  
 Db 1855 GACCAGAACACCCCGGGCTGTGCTGGTCCCGAGTAGCCTCCGATTTGAATCATATCCC 1914  
 QY 263 AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer 282  
 Db 1915 GCAGGGTCAACGGACTGCTGTGTACTTCCACCGTGGGCGACGACGACGAGAAACCTCG 1974  
 QY 283 AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisAsn 302  
 Db 1975 GCGCTGCATATCGTTGTTGGGAGCTCGCTGGCCATGGATGTCTCTAGTCCACCACAC 2034  
 QY 303 SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIle 322  
 Db 2035 AGCACACTCTTCTGCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2094  
 QY 323 LysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys 342

```
Db 2095 AAGGACAGTGAGAGAAACGGTGGTGGTCTTGCCAGATACGACTTTTCAGGTTTAAAG 2154
Qy 343 ThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGluHisThr 362
Db 2155 ACCTTCTCTCTCCCAACCATGCTATGGAGGGACAGTGTCTCTTCCCTCGCAACAACACAG 2214
Qy 363 ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer 382
Db 2215 GTGGGATCTCCAGGGATAGGAGCCCTGCCGGCAGGATGCTTTGTTGTCAGGCMAAGC 2274
Qy 383 LysGlnGlnLeuGluGluGlnGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp 402
Db 2275 AAGCAGCAGCTGGAGGAGCAGAGAAAGACACTGTATGTGTTTGGAAAGCTGCGCAGGAC 2334
Qy 403 ValGluGluTyrGlnValValCysGlyLysPheLeuAlaIleAenAlaThrAenMetSer 422
Db 2335 GTGAGGAGTGGCAAGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAAATGTGTC 2394
Qy 423 CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer 442
Db 2395 TGTGCTTGTCCCGGAGCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCT 2454
Qy 443 SerAspLeuIleLeuIleArgLysCysSerArgPheAenPheLeuArgPheLeuIleArg 462
Db 2455 TCTGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGG 2514
Qy 463 HisThrAenGlnAenAspGlnPheAspPheThrPheValGluValTyrArgValLysLys 482
Db 2515 CACACCAACCCAGCAGGACCACTTTGACTTCACTTTTGTGAAGTTTATCGCGTCAAGAAA 2574
Qy 483 PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLys 502
Db 2575 TTCCAGTTAGTTCGAGCACATGGAGGATGAGGACAGCAGCTCAAGAGGGGGGGAAG 2634
Qy 503 LysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSerAenSer 522
Db 2635 AAGCGCTTGGGCACATTTGCAGCAGCCACCCCTCTGCTGCTGCACCGTCTCCAACAGC 2694
Qy 523 SerTrpAenCysAspGlyValLeuHisSerProAlaIleGluValAtqValHisCys 542
Db 2695 TCCTGGAACTCGCAGCGGGGAGGTCTGCACAGCCCTGCCATCGAGGTCCAGAGTCCACTGC 2754
Qy 543 GlnLeuValArgLeuPheAlaArgGlyIleGluGluAenProLysProAspSerHisSer 562
Db 2755 CAGCTGGTTCGACTCTTTGCAGGGAATTGAGAGAAATCCGAAGCCAGACTCACACAGC 2814
```

## RESULT 7

```
US-09-969-896-9
; Sequence 9, Application US/09969896
; Publication No. US20030125533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1614
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-9
```

```
Alignment Scores: 6.05e-309 Length: 1614
Pred. No.: 2888.00 Matches: 537
Score:
```

```
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 95.47% Indels: 0
DB: 10 Gaps: 0
US-10-631-958-11 (1-562) x US-09-969-896-9 (1-1614)
Qy 26 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnArg 45
Db 1 ATGGGGCGGACGGGGCGGGCGGAGCCGCTGCAATCCGCTGCTGTGGGTGAAGCAGCAGCGC 60
Qy 46 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 65
Db 61 TGGCCCGCTGAGCCCTGGAGCCCGCGGGGCTCTGCTGGCTGGTGGCGGAGCCCGGGGCC 120
Qy 66 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal 85
Db 121 GGAGCCGGCCCCCGGGCGGATGCTCTCTGCTGCTGATCTGAGATCATCGCGTT 180
Qy 86 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 105
Db 181 GAGGAAACAGACGTTTCACGGGAAACATCAAGGCAGTGGAAAAATGGCAGAAAAATGGA 240
Qy 106 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAla 125
Db 241 CCTTACGCTTTTACAGTTCACTGTGTAAGAGACGACGCGCAACCGCTGGAAGTGGGCG 300
Qy 126 GlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 145
Db 301 CAGGTGACTTTCTCGTGTCCAGAGGAGCAGCTGTGTCTGTGGCTGACACCTCGCG 360
Qy 146 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAenProPhe 165
Db 361 GAGATGCTGGAGAGCTGACGCTCCAGACCAAGCATTTACTGTGTTATTTATCAACCCGTT 420
Qy 166 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 185
Db 421 GGAGAAAGACAAAGCAAGCGGATATATGAAGAAAGTGGCACCACCTGTTCCACCTTA 480
Qy 186 AlaSerIleThrThrAspIleIleValThrGluHisAlaAenGlnAlaLysGluThrLeu 205
Db 481 GCCTCCATACCACTGACATCATGTTACTGAACATGCTTAATCAGGCCAAGGACACTG 540
Qy 206 TyrGluIleAenIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 225
Db 541 TATGAGATTAACATAGACAAATACGACGGCATCTCTGTGTCGGCGGAGATGGTATGTTTC 600
Qy 226 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAen 245
Db 601 AGCGAGGTGCTGCACGGTCTGATTGGGAGGACGCGAGGAGCGCGGGGTCCGACAGAAC 660
Qy 246 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 265
Db 661 CACCCCGGGCTGTGCTGGTCCCCAGTAGCTCCCGGATGGAAATCATTTCCCGCAGGTC 720
Qy 266 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 285
Db 721 ACGGACTGCGTGTGTTACTCCACCGTGGGCACCGACGACGACAGAAACCTCGCGCTGCAT 780
Qy 286 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAenSerThrLeu 305
Db 781 ATCGTTGTTGGGACTCGCTGGCCATGGATGTGCTCTCAGTCCACCAACACAGCAGACTC 840
Qy 306 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer 325
Db 841 CTTGCTACTCCGTTGCTCGCTGGCTACGGCTTCTACGGGGACATCATCAAGGACAGT 900
Qy 326 GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 345
Db 901 GAGAAAGAAACGGTGGTGGGTCTTTGCCAGATACACACTTTTTCAGGTTTAAAGACCTTC 960
Qy 346 SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 365
Db 1
```

Db 961 TCCACCACCTGCTATGAAGGACAGTGTCTTCTCCCTGTCACACACACAGGTGGATCT 1020  
QY 366 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 385  
Db 1021 CCAAGGGATAGGAAGCCCTGCGGGCAGCATGCTTTGTTGCGAGCAAGCAGCAG 1080  
QY 386 LeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 405  
Db 1081 CTGGAGGAGGACAGAAAGAACACTGTATGTGTTGGAAGCTGCGAGGACGTGGAGGAG 1140  
QY 406 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 425  
Db 1141 TGGCAAGTGTCTGTGGGAAGTTCTGGCCATCAATGCCACAAACATGTCCTGTCTGT 1200  
QY 426 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 445  
Db 1201 CGCCGAGGCCCGAGGGCCCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCGACCTC 1260  
QY 446 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 465  
Db 1261 ATCCCTCATCCGGAATGCTCCAGGTTCATATTTCTGAGATTTCATCAAGGCACACCAAC 1320  
QY 466 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 485  
Db 1321 CAGCAGGACCACTTGTGACTTCACTTTGTTGAAGTTTATCGGTCAAGAAATTCAGTTT 1380  
QY 486 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe 505  
Db 1381 ACCTCGAAGCACATGAGGATGAGACAGCACCTCAAGCAGGGGGGAAGAGCGCTTT 1440  
QY 506 GlyHisIleCysSerSerHisProSerCysCysCysThrValSerAsnSerSerTrpAsn 525  
Db 1441 GGGCACATTTGACGACGACCCCTCTCTGCTGCTGACCGTCTCAACACAGCTCCCTGGAAC 1500  
QY 526 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 545  
Db 1501 TCGACGGGAGGTCTGCACAGCCCTGCCATCGAGGTCAAGTCCAGTCCAGCTGGTT 1560  
QY 546 ArgLeuPheAlaArgGlyIleGluLeuAsnProLysProAspSerHisSer 562  
Db 1561 CGACTCTTTGCACGAGGAATTGAAGAAATCCGAAGATCCGAAGCAGACTCACACAGC 1611

## RESULT 8

US-10-631-958-9

; Sequence 9, Application US/10631958

; Publication No. US20040192580A1

; GENERAL INFORMATION:

; APPLICANT: Kossida, Sophia

; TITLE OF INVENTION: Regulation of human Sphingosine

; FILE OF INVENTION: Kinase-Like Protein

; FILE REFERENCE: 004974.00594

; CURRENT APPLICATION NUMBER: US/10/631,958

; PRIORITY FILING DATE: 2003-08-01

; PRIORITY APPLICATION NUMBER: US/09/969,896

; PRIORITY FILING DATE: 2001-10-04

; PRIORITY APPLICATION NUMBER: US 60/238,005

; PRIORITY FILING DATE: 2000-10-06

; PRIORITY APPLICATION NUMBER: US 60/314,113

; PRIORITY FILING DATE: 2001-08-23

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 9

; LENGTH: 1614

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-631-958-9

Alignment Scores:

Pred. No.: 6.05e-309 Length: 1614

Score: 2888.00 Matches: 537

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 95.47% Indels: 0

DB: 19 Gaps: 0

US-10-631-958-11 (1-562) x US-10-631-958-9 (1-1614)

QY 26 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnGlnArg 45  
Db 1 ATGGGGGGCGACGGGGGGCGGGAGCGCTGCANTCCGTGCTGTGGGTGNAGCAGCAGCGC 60  
QY 46 CysAlaValSerLeuGluProAlaAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 65  
Db 61 TCGCCGCTGAGCTGAGCGCGCGCGCTCTGCTGCGCTGCTGCGCGAGCCCGGGCGCC 120  
QY 66 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleIleAlaVal 85  
Db 121 GGAGCCGGCGCCCGCGCGGATGCCCTGCTCTGCTGCTGCTATCTGAGATCATCCCGCTT 180  
QY 86 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 105  
Db 181 GAGGAAACAGACGTTACGGGAAACATCAAGCAGTGGGAAATGCGAGNAATGGGAAAG 240  
QY 106 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAla 125  
Db 241 CCTTACGCTTTTACAGTTTCACTGTGTAAGAGAGACGACGCGCCTGGAAGTGGGCG 300  
QY 126 GlnValThrPheTrpCysProGluGlnLysCysHisLeuTrpLeuGlnThrLeuArg 145  
Db 301 CAGGTGACTTTTCTGGTGTCCAGAGGAGCAGTGTGTCTCTTGTGCTGCAGACCTGCGG 360  
QY 146 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 165  
Db 361 GAGATGCTGGAGAGCTGACGTCCAGACCAAGCATTTTCTGTTATTTATCAACCCGTTT 420  
QY 166 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 185  
Db 421 GGAGAAAGACACAAGGCAAGCGGATATGAAAGAAAGTGGCACCACTGTTTCACTTCA 480  
QY 186 AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 205  
Db 481 GCCTCCATCCACCTGACATCATCTGTTACTGAACATGCTTAATCAGGCCCAAGGAGACT 540  
QY 206 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 225  
Db 541 TATGAGATTACATAGACAAATACAGCGCATCTGCTGTCTGCTGCGCGGAGATGTTATGTC 600  
QY 226 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 245  
Db 601 AGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACGAGAGCGCGCGGGTCCACCAAGAC 660  
QY 246 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProIleGlySer 265  
Db 661 CACCCCGGGCTGTGTGCTGCCAGTAGCTCCCGGATTTGGAATCATTTCCCGCAGGGTCA 720  
QY 266 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 285  
Db 721 ACAGACTGCTGTGTACTCCACCCTGGGACACGACGACGACGACCAAACTCCGCGCTGCAT 780  
QY 286 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 305  
Db 781 ATCGTGTGTTGGGACTCGCTGGCCATGATGTCTCAGTCCACCAACCAACAGCACATC 840  
QY 306 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer 325  
Db 841 CTTCGCTACTCCGTGCTCCCTGCGGCTTCTACGGGGAATCATCATCAAGGACAGT 900  
QY 326 GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 345  
Db 901 GAGAGAAACCGTGTGGGTCTTGCCAGATACGACTTTTCAAGTTTAAAGACCTTCCTC 960  
QY 346 SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 365  
Db 961 TCCACCACTGCTATGAAGGACAGTGTCTCTCTCCCTGTCACACACACACCGTGGATCT 1020  
QY 366 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 385



```
Db 1021 CCAAGGATAGAAAGCCCTGCGGCGAGATGCTTTGTTTTCAGGCAAGCAAGCAGCAG 1080
Qy LeuGluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAAspValGluGlu 405
Db 1081 CTGAGGAGGAGCAGAGAAAGCACTGATGTTGTTGGAAGCTGCGAGGAGCTGGAGGAG 1140
Qy TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 425
Db 1141 TGGCAAGTCGTCGTGGGAAGTTCTGGCCATCAATGCCACAACATGTCCTGTGCTGT 1200
Qy ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 445
Db 1201 CGCGGAGCCCAAGGGCTCTCCCGGCTGCCACTTGGGAGACGGGCTCTTCTGACCTC 1260
Qy IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 465
Db 1261 ATCTCATCCGGAATGCTCCAGGTTCAAATTTCTGAGATTTCTCATCAGGACACCAAC 1320
Qy GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 485
Db 1321 CAGCAGGACCAAGTTGACTTCACTTTGTGGAAGTTTATCGGTCAGAAATTCAGTTT 1380
Qy ThrSerLysHisMetGluAAspGluAspSerAspLeuLysGluGlyLysLysArgPhe 505
Db 1381 ACGTCGAAGCACATGGAGGATCAGACAGCAGCCTCAAGGAGGGGGGAAGCGCTTT 1440
Qy GlyHisLysCysSerSerHisProSerCysCysCysThrValSerAsnSerSerTrpAsn 525
Db 1441 GGGCACATTTGAGCAGCAGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
Qy CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 545
Db 1501 TGGCAGCGGGAGGCTCTGCACAGCCCTGCCATCGAGGTCAGAGTCCAGTCCAGCTG 1560
Qy ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 562
Db 1561 CGACTCTTTGACAGGGAATTTGAAGAAATCCGAAGCCAGACTCACACAGC 1611

RESULT 9
US-09-784-810A-5
; Sequence 5, Application US/09784810A
; Patent No. US20020082203A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: SAME
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/09/784,810A
; CURRENT FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1840
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-784-810A-5

Alignment Scores:
Pred. No.: 1,9e-282 Length: 1840
Score: 2649.50 Matches: 495
Percent Similarity: 97.83% Conservative: 1
Best Local Similarity: 97.63% Mismatches: 0
Query Match: 87.59% Indels: 11
DB: 9 Gaps: 1

US-10-631-958-11 (1-562) x US-09-784-810A-5 (1-1840)
Qy 67 AlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleIleAlaValGlu 86
```

```
Db 1 GCCGCGCCCGCGCGGAGATGCCCTGCTCTGTGCTGTATCTGAGATCATCGCGCTTGAG 60
Qy GluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLysPro 106
Db 61 GAAACACAGCTTCACGGGAACATCAAGGCAAGTGGAAAATGGCAGAAAATGGAAAAGCCT 120
Qy TyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAlaGln 126
Db 121 TAGCTTTTACAGTTCACTGTGTAAGAGAGACGACGCGCACCGCTGGAAGTGGCGCAG 180
Qy ValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArgGlu 146
Db 181 GTGACTTTCTGGTCTCCAGAGGAGCAGCTGTGTCACTTGTGGCTGACAGACCTCGCGGAG 240
Qy MetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPheGly 166
Db 241 ATGCTGGAGAGCTGAGCTCCAGACCAAGCAATTTACTTGGTATTTATCAACCCCTTTGGA 300
Qy GlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeuAla 186
Db 301 GGAAGAGACAGGCAAGCGGATATATGAAAGAAAGTGGCACCACCTGTTCCACCTTAGCC 360
Qy SerIleThrThrAspIleIle-----ValThr 195
Db 361 TCCATCACTCATCATCGGTAAACAAATTCCTATGTTAACTATGTAGAAGTAATTA 420
Qy GluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTrpAspGly 215
Db 421 GAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTTAACTAGACAAATACAGCGC 480
Qy IleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLysGlyArg 235
Db 481 ATCGTCTGTGCGCGGAGATGGTATGTTTCAGCGAGGTGCTGCACGCTCTGATTGGGAGG 540
Qy ThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSer 255
Db 541 ACGCAGAGGAGCGCGGGGTCGACACAGAACACCCCGGGGCTGTGCTGGTCCCAGTAGC 600
Qy LeuArgIleGlyIleIleProAlaGlySerThrAspCysValCysTyrSerThrValGly 275
Db 601 CTCGGATTTGGATCATTTCCGCGAGGTTCAACGAGACTGCGTGTGTTACTCCACCGTGGGC 660
Qy ThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAsp 295
Db 661 ACCAGCGACGAGAAACCTCGCGGCTGCATATCGTTGTTGGGACTCGCTGGCCATGGAT 720
Qy ValSerSerValHisHisAsnSerThrIleuArgTyrSerValSerLeuLeuGlyTyr 315
Db 721 GTGTCTCTCAGTCCACCAACACAGCACACTCTCTCGCTACTCCGTGTCCTCTGGGCTAC 780
Qy GlyPheTyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuLysLeuAlaArg 335
Db 781 GGCCTTCTACGGGGACATCATCAGGACAGTGAAGAAACCGTGGTGGTCTTGGCCAGA 840
Qy TyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSer 355
Db 841 TACGACTTTTACGGTTTAAAGACCTTCTCTCTCCACCACTGCTATGAAGGACAGTGTCC 900
Qy PheLeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGly 375
Db 901 TTCTCTCCCTGCACAAACACACCGTGGATCTCTCAAGGGATAGGAAGCCCTGCGCGGACGA 960
Qy CysPheValCysArgGlnSerLysGlnLeuGluGlnLysLysAlaLeuTyr 395
Db 961 TGCCTTTGTTTCAGCAGCAAGACACAGCAGCTGGAGGAGGAGCAAGAAAGCACTGTAT 1020
Qy GlyLeuGluAlaAlaGluAspValGluIleTrpGlnValValCysGlyLysPheLeuAla 415
Db 1021 GGTTTGGAAGCTGCGGAGGACGTGGAGGTGGCAAGTCTGCTGTGGGAAGTTTCTGCGCC 1080
Qy IleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAla 435
```



D	b		1081	ATCAATGCCACAAACATGTCCTGTGGTCTTGTCGCGAGAGCCCCAGGGGCCTTCCTCCCGGCT	1141
Q	y		436	AlaHisLeuGlyAspGlySerAspLeulleLeuilleAargLysCysSerArgPheAsn	455
D	b		1141	GCCCACTTGGAGACGGGTCTTTCTGACTCATCTCATCCGAAAATGCTCCAGGTTCAAT	1200
Q	y		456	PheLeuArgPheLeulleAargHisThrAsnGlnAspGlnPheAspPheThrPheVal	475
D	b		1201	TTCCTGAGATTCTCATCAGCACACCAGCAGGACCATTTTGACTTCACTTTTGT	1260
Q	y		476	GluValTyArgValLysLysPheGlnPheThrSeryIshMetGluAspGluAspSer	495
D	b		1261	GAAGTTTATCCGCTCAAGAANTCCAGTTTACGTCTGAAGCACATGGAGGATGAGGACGC	1320
Q	y		496	AspLeuLysGluGlyLysLysAargPheGlyHisileCysSerSerHisProSerCys	515
D	b		1321	GACCTCAAGAGGGGGGAAGACGGCTTTGGGCACATTTGCAGCAGCCACCCCTCTGC	1380
Q	y		516	CysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAla	535
D	b		1381	TGCTGCACCGCTCTCCAACAGCTCTGGAACTCGACCGGGAGGTCTTGCACAGCCCTGCC	1440
Q	y		536	IleGluValArgValHisCysGlnLeuValArgLeuPhealaargGlyIleGluGluAsn	555
D	b		1441	ATCCAGGTCAAGGTCCACTGCCAGCTGGTTCGACTCTTTGCAGGAGAAATTGAAGAGAT	1500
Q	y		556	ProLysProAspSerHisSer	562
D	b		1501	CCGNAGCCAGNCTCACAGC	1521
<b>RESULT 10</b>					
US-10-876-281-5					
; Sequence 5, Application US/10876281					
; Publication No. US20050123942A1					
; GENERAL INFORMATION:					
; APPLICANT: RASTELLI, LUCA					
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING					
; FILE REFERENCE: SAME					
; CURRENT APPLICATION NUMBER: US/10/876,281					
; CURRENT FILING DATE: 2004-06-24					
; PRIOR APPLICATION NUMBER: US/09/784,810					
; PRIOR FILING DATE: 2001-02-14					
; PRIOR APPLICATION NUMBER: 60/182,360					
; PRIOR FILING DATE: 2000-02-14					
; PRIOR APPLICATION NUMBER: 60/191,261					
; PRIOR FILING DATE: 2000-03-22					
; NUMBER OF SEQ ID NOS: 29					
; SOFTWARE: PatentIn Ver. 2.1					
; SEQ ID NO 5					
; LENGTH: 1840					
; TYPE: DNA					
; ORGANISM: Homo sapiens					
US-10-876-281-5					
Alignment Scores:					
Pred. No.:            1,9e+282          Length:         1840					
Score:                2649.50          Matches:         495					
Percent Similarity:   97.83%          Conservative:    1					
Best Local Similarity: 97.63%          Mismatches:    0					
Query Match:           87.59%          Indels:          11					
DB:                      22               Gaps:            1					
US-10-631-958-11 (1-562) x US-10-876-281-5 (1-1840)					
Q	y		67	AlaGlyAlaproGlyAlaAspAlaCysSerValProValSerGluIlelleAlaValGlu	86
D	b		1	GCCGGCGCCCCCGCGCGGATGCTCTGTGCTCTTAICTGAGATCATCGCGTTGAG	60
Q	y		87	GluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLysPro	106
D	b		61	GAATAACAGCGTTTCACGGGMAAACATCAAGGCGATGGAAAAATGGCAGAAAAATCGAAAGCCT	120

```
Db 1201 TTTCTGAGATTCTCATCAGGCACACCAACAGCAGGACCAGTTGTGACTTTCACCTTTTGTT 1260
Qy 476 GluValTyrArgValLysPheGlnPheThrSerLysHisMetGluAspGluAspSer 495
Db 1261 GAAGTTTATCGCGTCAAGAAATTCAGTTTACGTGCGAAGCAATGGAGGATGAGGACAGC 1320
Qy 496 AspLeuLysGluGlyGlyLysLysArgPheGlyHisLysCysSerSerHisProSerCys 515
Db 1321 GACCTCAAGAGGGGGGAAGAAGCGCTTGGGACACATTTGCACAGCCACCCCTCTGCG 1380
Qy 516 CysCysThrValSerAsnSerSerTriaAsnCysAspGlyGluValLeuHisSerProAla 535
Db 1381 TGCTGCACCGCTCCAAACAGCTCTCGAATGCGACGGGAGGTCTCTGCACAGCCCTGCC 1440
Qy 536 IleGluValArgValHisCysGluValArgLeuPheAlaAargGlyIleGluGluAsn 555
Db 1441 ATCAGAGTCAAGGGTCCACTGCCAGCTGGTTCCGACTCTTTGCACGAGGAATTGAAGAAAT 1500
Qy 556 ProLysProAspSerHisSer 562
Db 1501 CCGAAGCCAGACTCACACAGC 1521

RESULT 11
US-09-969-896-1
; Sequence 1, Application US/09969896
; Publication No. US20030125533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969.896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 979
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-1

Alignment Scores:
Pred. No.: 4.17e-171 Length: 979
Score: 1640.50 Matches: 314
Percent Similarity: 96.63% Conservative: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 54.23% Indels: 11
DB: 1.0 Gaps: 1

US-10-631-958-11 (1-562) x US-09-969-896-1 (1-979)
Qy 155 ProLysHisLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 174
Db 2 CCAAGCAATTTACTGTGTTATTAATCAACCGTTTGAGGAAAGGACAGGCGCAGGATA 61
Qy 175 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIle--- 193
Db 62 TATGAAGAAAGTGGCACCACCTGTTCACTTAGCTCCATCACCACCTGACATCATCGGT 121
Qy 194 -----ValThrGluHisAlaAsnGlnAlaLysGlu 203
Db 122 AACAAATTTCTATGTTAACTATGTAGAAGTAATTAATCAACATGCTAATCAGGCCAAGGAG 181
Qy 204 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 223
Db 182 ACTCTGTATGAGATTAAATAGAAATACACGCGCATCGTGTGTGCGGGAGATGGT 241
Qy 224 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 243
```

```
Db 242 ATGTTTCAGCCAGGTGCTGCACCGGTCTGATTGGGAGGACGACGAGGAGCGCGGGCTCGAC 301
Qy 244 GluAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 263
Db 302 CAGAACACACCCCGGGGTGTGTGGTCCCGAGTAGCTCCGGATTGGAAATCAITTCGCCGA 361
Qy 264 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 283
Db 362 GGGTCAACGAGACTGCGGTGTGTACTCCACCGTGGGACACGACGACGACGAGAAACCTCGCGC 421
Qy 284 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 303
Db 422 CTGCATATCGTTGTGGGACTCGCTGGCCATGTGTCTCAGTCCAGTCCACCAACACAGC 481
Qy 304 ThrLeuLeuArgTyrSerValSerLeuGlyTyrGlyPheTyrGlyAspIleIleLys 323
Db 482 ACACCTCTTCGCTACTCCGTGTCCCTGTGGGTCTTACCGGGGACATCATCAAG 541
Qy 324 AspSerGluLysLysArgTyrLeuAlaArgTyrAspPheSerGlyLeuLysThr 343
Db 542 GACAGTGAGAAGAAACGCTGGTGGGTCTTCCAGATACGACTTTTCAGGTTTAAAGACC 601
Qy 344 PheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 363
Db 602 TTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACACACACCGGTG 661
Qy 364 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 383
Db 662 GSATCTCCAAAGGATAGAAAGCCCTCCCGGGAGGATGCTTTGTTTCAGGCAAGCAAG 721
Qy 384 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 403
Db 722 CAGCAGCTGGAGGAGGAGCAGAGAAGACACTGTATGTTTGGAGCTGCGGAGACGTG 781
Qy 404 GluGlnTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 423
Db 782 GAGGAGTGGCAAGTCTGTGGGAAGTTTCTGCCCATCAATGCCACAAACATGCTCTGT 841
Qy 424 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer 443
Db 842 GCTTGTGCGCGAGCCCGGAGGCTCTCCCGGCTGCCCATTTCTGAGATTTCATCAGGCAC 901
Qy 444 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 463
Db 902 GACCTCATCTCATCCGGAATGCTCCAGGTTCATTTTCTGAGATTTCATCAGGCAC 961
Qy 464 ThrAsnGlnGlnAspGln 469
Db 962 ACCAACCCAGCAGGACCAG 979

RESULT 12
US-10-631-958-1
; Sequence 1, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631.958
; PRIOR FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 979
; TYPE: DNA
```

; ORGANISM: Homo sapiens			
US-10-631-958-1			
Alignment Scores:			
Pred. No.:	4.17e-171	Length:	979
Score:	1640.50	Matches:	314
Percent Similarity:	96.63%	Conservative:	1
Best Local Similarity:	96.32%	Mismatches:	0
Query Match:	54.23%	Indels:	11
DB:	19	Gaps:	1
US-10-631-958-11 (1-562) x US-10-631-958-1 (1-979)			
Qy	155	ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyAsnArgile	174
Db	2	CCAAAGCATTTACTGGTATTTATCAACCGTTTGAGGAGAAAGGCAAGGCAAGCGGATA	61
Qy	175	TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIle---	193
Db	62	TATGAAGAAGAAAGTGGACCACTGTTACCTTAGCTTCATCCACCTGACATCATTCGGT	121
Qy	194	-----ValThrGluHisAlaAsnGlnAlaLysGlu	203
Db	122	AACAAATCTATGTTAACTATGTAGAAGTAATTACTGAACATGCTAATCAGGCCAAGGAG	181
Qy	204	ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly	223
Db	182	ACTCTGTATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGTGGGGAGATGGT	241
Qy	224	MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp	243
Db	242	ATGTTACAGGAGTCTGCACGGTCTGATTTGGAGAGCGAGAGGCGCGGGTGCAC	301
Qy	244	GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla	263
Db	302	CAGAACCAACCCCGGGCTGTGCTGCCAGTAGCTCCGGATTGGAATCATTTCCGCA	361
Qy	264	GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla	283
Db	362	GGGTCAACCGGACTGGCTGTGTACTCCACCGTGGGCACGACGACAGAAACCTCGGCG	421
Qy	284	LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer	303
Db	422	CTGCATATCGTTGTGGGACCTCGCTGGCCATGGATGTCTCTAGTCCACCAACACAGC	481
Qy	304	ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys	323
Db	482	ACACTCCTTCGCTACTCCGCTGTCCCTGTGGCTACGGCTTCTACGGGGACATCATCAAG	541
Qy	324	AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr	343
Db	542	GACAGTGAGAGAAACCGGTGGTGTGGTCTTGCCAGATACGACTTTTTCAGGTTTAAAGACC	601
Qy	344	PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal	363
Db	602	TTCTCTCTCCCACTCTGTATGAAGGAGCAGTGTCTTCTTCCCTTCGACACACACCGGTG	661
Qy	364	GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys	383
Db	662	GGATCTCCAGGGATAGGAAGCCCTGCCCGGACGAGATGCTTTGTTTGAGGCNAAGCAAG	721
Qy	384	GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal	403
Db	722	CAGCAGCTGGAGGAGGAGCAGAAAGCACTGTATGGTTTGGAGCTGCGGAGCAGCTG	781
Qy	404	GluGluTyrGlnValValCysGlyLysPheIleuAlaIleAsnAlaThrAsnMetSerCys	423
Db	782	GAGGAGTGGCAAGTCTGTGGGGAAGTTTCTGGCCATCAATGCCACAAAATGTCTCTGT	841
Qy	424	AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer	443
Db	842	GCTTGTCCCGGAGGCCCCAGSGGGCTCTCCCGGGCTGCCCATCTTGGGAGACGGGTCTTCT	901

---

Qy	444	AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis	463
Db	902	GACCTCATCTCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTCTCATCGGCAC	961
Qy	464	ThrAsnGlnGlnAspGln	469
Db	962	ACCAACCAGCAGGACCAG	979
RESULT 13			
US-10-115-635-247			
; Sequence 247, Application US/10115635			
; Publication No. US20040137434A1			
; GENERAL INFORMATION:			
; APPLICANT: Tang, Y. Tom			
; APPLICANT: Zhou, Ping			
; APPLICANT: Goodrich, Ryle			
; APPLICANT: Liu, Chenghua			
; APPLICANT: Asundi, Vinod			
; APPLICANT: Ren, Feiyan			
; APPLICANT: Zhang, Jie			
; APPLICANT: Zhao, Qing A.			
; APPLICANT: Xue, Aidong J.			
; APPLICANT: Yang, Yonghong			
; APPLICANT: Wehrman, Tom			
; APPLICANT: Drmanac, Radoje T.			
; TITLE OF INVENTION: Novel Nucleic Acids and			
; FILE REFERENCE: Polypeptides			
; CURRENT APPLICATION NUMBER: US/10/115,635			
; CURRENT FILING DATE: 2002-04-03			
; PRIOR APPLICATION NUMBER: 09/714,936			
; PRIOR FILING DATE: 2000-11-17			
; NUMBER OF SEQ ID NOS: 362			
; SOFTWARE: pt FL_genes Version 2.0			
; SEQ ID NO 247			
; LENGTH: 817			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (90)..(662)			
US-10-115-635-247			
Alignment Scores:			
Pred. No.:	1.45e-117	Length:	817
Score:	1156.00	Matches:	213
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	38.21%	Indels:	0
DB:	19	Gaps:	0
US-10-631-958-11 (1-562) x US-10-115-635-247 (1-817)			
Qy	3	AlaAlaAsnGlyProAlaProLeuGlyValArgAlaProProAlaTyrArgThrSerPro	22
Db	21	GCCGCTAACGGTCCGCGCGCCCTCCGCGTCCGCGCGCCCGCCAGCTGGCGACGAGCCG	80
Qy	23	AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTyrValLys	42
Db	81	CGGCGGAGATGGGGGCGACGGGGGCGCGGAGCCGTGCANTCCGTGCTGTGGGTGAG	140
Qy	43	GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTyrTyrArgSer	62
Db	141	CAGCAGCGCTCGCGCGTGAGCTGGAGCCCGCGCGGCTCTGCTGCGCTGGTGGCGAGC	200
Qy	63	ProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIle	82
Db	201	CCGGGCGCGGAGCCGCGCGCCCGCGCGGATGCCCTGCTGTGCTGTATCTTGAGATC	260
Qy	83	IleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTyrGlnLys	102
Db	261	ATCCCGCTTGAGGNAACAGAGCTTCACGGGNAACATCAAGGCAGTGGNAATGCGAGNA	320

```
Qy 103 MetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTyr 122
Db 321 ATGGAAGAGCTTACGCTTTTACAGTTTACAGTTTGTAAAGAGAGACGACGCGCTGG 380
Qy 123 LysTrpAlaGlnValThrPheTyrCysProGluGlnLeuCysHisLeuTyrProGln 142
Db 381 AAGTGGCGCAGGTGACTTCTTGCTGTCCAGAGGAGCAGCTGTGTCACTTGTGGCTGCAG 440
Qy 143 ThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIle 162
Db 441 ACCCTGGGGAGATGCTGGAGAGCTGACGTCCAGACCAAGCAATTTACTGGTATTATC 500
Qy 163 AsnProPheGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeu 182
Db 501 AACCCGTTTGGAGAAAGAGCAAGGCAAGCGGATATATGAAAGAAAGTGGCACCACTG 560
Qy 183 PheThrLeuAlaSerIleThrAspIleIleValThrGluHisAlaAsnGlnAlaLys 202
Db 561 TTCACCTTAGCTCCATCACCACCTGACATCATCGTTACTGAACATGCTTAATCAGGCCAAG 620
Qy 203 GluThrLeuTyrGluIleAsnIleAspLysTyrAspGly 215
Db 621 GAGACTCTGTATGAGATTAACTAGACAAATACGACGGG 659
```

## RESULT 14

```
US-09-969-896-8
; Sequence 8, Application US/09969896
; Publication No. US20030125533A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 8
```

```
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-8
```

```
Alignment Scores:
Pred. No.: 2e-101 Length: 550
Score: 1008.00 Matches: 183
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 33.32% Indels: 0
DB: 10 Gaps: 0
```

```
US-10-631-958-11 (1-562) x US-09-969-896-8 (1-550)
```

```
Qy 1 HisGluAlaAlaAsnGlyProAlaProLeuGlyValArgAlaProProAlaTyrArgThr 20
Db 1 CACGAGGCGGTAAACGGTCCGCGCCCCCTCGCGCTCCGCGCGCCCCCAGCGCTGGCGGACG 60
Qy 21 SerProAlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTyr 40
Db 61 AGCCCGGCGGGAGATGGGGCGGAGCGGGCGGCGCGCGCTGCATCCGTCTGTGG 120
Qy 41 ValLysGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTyrTrp 60
Db 121 GTGAAGCAGCAGCGCTGCGCGCTGAGCTGAGCGCGCGCGGCTCTGCTGCGCTGGTGG 180
Qy 61 ArgSerProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSer 80
Db 181 CGGAGCCCGGGCGCGGAGCCCGCGCGCGGATGCTGCTGTGCTGCTATCT 240
```

```
Qy 81 GluIleIleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTyr 100
Db 241 GAGATCATCGCGCTTGAGGAAAACAGACGTTTACGGGAAAACATCAAGGACGTGGAATAATGG 300
Qy 101 GlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgHis 120
Db 301 CAGAAAATGGAAGCCCTTACGCTTTTACAGTTTCACTGTGTAAAGAGAGACGACGCGCAC 360
Qy 121 ArgTrpLysTrpAlaGlnValThrPheTyrCysProGluGlnLeuCysHisLeuTyr 140
Db 361 CGCTGGAGTGGGCGCAGGTGACTTCTGTGTCCAGAGGAGCAGCTGTGTCACTTGTGG 420
Qy 141 LeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuVal 160
Db 421 CTGCAGACCTTGGCGGAGATGCTGGAGAGCTGACGTCCAGACCAAGAGCATTTACTGGTA 480
Qy 161 PheIleAsnProPheGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAla 180
Db 481 TTTATCAACCCGTTTGGAGGAAAAGGACAAAGCGGATATATGAAAGAAAAGTGGCA 540
Qy 181 ProLeuPhe 183
Db 541 CCACCTGTTT 549
```

## RESULT 15

```
US-10-631-958-8
; Sequence 8, Application US/10631958
; Publication No. US20040192580A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 8
```

```
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-958-8
```

```
Alignment Scores:
Pred. No.: 2e-101 Length: 550
Score: 1008.00 Matches: 183
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 33.32% Indels: 0
DB: 19 Gaps: 0
```

```
US-10-631-958-11 (1-562) x US-10-631-958-8 (1-550)
```

```
Qy 1 HisGluAlaAlaAsnGlyProAlaProLeuGlyValArgAlaProProAlaTyrArgThr 20
Db 1 CACGAGGCGGTAAACGGTCCGCGCCCCCTCGCGCTCCGCGCGCCCCCAGCGCTGGCGGACG 60
Qy 21 SerProAlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTyr 40
Db 61 AGCCCGGCGGGAGATGGGGCGGAGCGGGCGGCGCGCGCTGCATCCGTCTGTGG 120
Qy 41 ValLysGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTyrTrp 60
Db 121 GTGAAGCAGCAGCGCTGCGCGCTGAGCTGAGCGCGCGCGGCTCTGCTGCGCTGGTGG 180
Qy 61 ArgSerProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSer 80
```

```
Db      181 CGGAGCCCGGGCCCGGAGCCGGCGCCCGCGCGGATGCTGCTGTGCTGTATCT 240
Qy      81  GluIleIleAlaValGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrp 100
Db      241 GAGATCATCGCGTTTGAGGAAACAGACGTTTCACGGGAACATCAAGGCAGTGGAAAATGG 300
Qy      101 GlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgArgHis 120
Db      301 CAGAAATGGAAAGCCTTACGCTTTTACAGTTCACTGTGTAAAGAGACGACGGCAC 360
Qy      121 ArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrp 140
Db      361 CGCTGGAAGTGGCGCAGGTGACTTCTGGTGTCCAGAGGAGCAGCTGTGTCACTTGTGG 420
Qy      141 LeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuVal 160
Db      421 CTGCAGACCCCTCGGGGAGATGCTGGAGAAGCTGACGTCAGACCAAGCATTTACTGGTA 480
Qy      161 PheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAla 180
Db      481 TTTATCAACCCGTTTGGAGGAAGGACAAAGCCAGCGGATATATGAAAGAAAAGTGGCA 540
Qy      181 ProLeuPhe 183
Db      541 CCACTGTTC 549
```

Search completed: September 6, 2005, 10:49:17  
Job time : 1646.07 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 6, 2005, 02:19:54 ; Search time 867.649 Seconds  
(without alignments)  
3834.378 Million cell updates/sec

Title: US-10-631-958-11

Perfect score: 3025

Sequence: 1 HEANGPAPLGVAPPWPT.....QLVRLFARGIENKPDPSHS 562

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US10631958/runat.02092005.165814.4497/app.query.fasta\_1.1941  
-DB=N\_Geneseq\_16Dec04 -QFMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10631958 @CGN.1.1436 @runat.02092005.165814.4497 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_16Dec04:\*

Result No.	Score	Query Match	Length	ID	Description
1	3025	100.0	1740	8	ADA05679
2	3025	100.0	4413	6	ABL40828
3	3012	99.6	4429	12	ADJ96598
4	3012	99.6	4445	13	ADP55247
5	3004	99.3	4463	6	ABA96945

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3025	100.0	1740	8	ADA05679
2	3025	100.0	4413	6	ABL40828
3	3012	99.6	4429	12	ADJ96598
4	3012	99.6	4445	13	ADP55247
5	3004	99.3	4463	6	ABA96945

6	2980	98.5	1740	12	ADNG2844
7	2931.5	96.9	4432	8	ABX70921
8	2649.5	87.6	1840	5	AAI14426
9	2587	85.5	4231	3	AAA50510
10	2541.5	84.0	4702	13	ADS10370
11	2400	79.3	2241	5	AA577728
12	2210	73.1	3975	8	ACC70838
13	1714.5	56.7	2186	5	AA577730
14	1640.5	54.2	979	6	ABL40822
15	1386.5	45.8	1570	5	AA577731
16	1156	38.2	817	6	ABQ99499
17	1156	38.2	817	13	ADS11669
18	1055	34.9	753	3	AA576031
19	1008	33.3	550	6	ABL40827
20	986	32.6	547	3	AA576592
21	811	26.8	474	6	ABL40823
22	721.5	23.9	10337	4	ABL25704
23	686.5	22.7	2173	4	ABL25705
24	679	22.4	426	5	AA577727
25	656	21.7	522	5	AA577727
26	626	20.7	411	5	AA577729
27	601	19.9	1774	12	ADQ88890
28	541.5	17.9	15185	4	AA575589
29	525	17.4	329	6	ABL40824
30	511.5	16.9	15181	4	AA575588
31	437.5	14.5	564	8	ABT23453
32	343.5	11.4	382	5	AA574428
33	331.5	11.0	2830	4	ABL02327
34	329.5	10.9	2609	9	ACF35860
35	310.5	10.3	2020	4	ABL03297
36	310.5	10.3	2629	9	ACF35859
37	310.5	10.3	4020	4	ABL03296
38	302	10.0	167	6	ABL40825
39	298.5	9.9	2422	5	AA585331
40	295.5	9.8	2682	13	ADQ83707
41	295.5	9.8	2682	12	ADQ87011
42	295.5	9.8	2875	5	ABX71285
43	288.5	9.5	1857	6	ABK90199
44	288.5	9.5	2380	5	AA514817
45	288.5	9.5	2380	6	ABL59533

ALIGNMENTS

RESULT 1

ADA05679

ID ADA05679 standard; cDNA; 1740 BP.

XX AC ADA05679;

XX DT 06-NOV-2003 (first entry)

XX XX Human NOV9a encoding cDNA SEQ ID NO:39.

human; NOV9a; antidiabetic; anorectic; antibacterial; virucide;  
immunomodulator; cytosolic; neurotropic; neuroprotective;  
antiparkinsonian; antilipemic; gene therapy; human disease;  
metabolic disorder; diabetes; obesity; infection; cachexia; cancer;  
neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
immune disorder; haematopoietic disorder; dyslipidaemia; gene; ss.

XX OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 76..1689

XX FT /\*tag= a

XX FT /product= "NOV9a"

XX PN WO2003029424-A2.

XX PD 10-APR-2003.

XX PF 02-OCT-2002; 2002WO-US031373.



XX PR 02-OCT-2001; 2001US-0326483P.  
XX PR 05-OCT-2001; 2001US-0327435P.  
XX PR 05-OCT-2001; 2001US-0327449P.  
XX PR 09-OCT-2001; 2001US-0327917P.  
XX PR 09-OCT-2001; 2001US-0328029P.  
XX PR 09-OCT-2001; 2001US-0328044P.  
XX PR 09-OCT-2001; 2001US-0328056P.  
XX PR 12-OCT-2001; 2001US-0328849P.  
XX PR 15-OCT-2001; 2001US-0329414P.  
XX PR 17-OCT-2001; 2001US-0330142P.  
XX PR 18-OCT-2001; 2001US-0330309P.  
XX PR 22-OCT-2001; 2001US-0341058P.  
XX PR 24-OCT-2001; 2001US-0339266P.  
XX PR 24-OCT-2001; 2001US-0343629P.  
XX PR 29-OCT-2001; 2001US-0349575P.  
XX PR 01-NOV-2001; 2001US-0346337P.  
XX PR 17-APR-2002; 2002US-0373260P.  
XX PR 19-APR-2002; 2002US-0373815P.  
XX PR 19-APR-2002; 2002US-0373817P.  
XX PR 19-APR-2002; 2002US-0373826P.  
XX PR 19-APR-2002; 2002US-0373884P.  
XX PR 22-APR-2002; 2002US-0374977P.  
XX PR 16-MAY-2002; 2002US-0381037P.  
XX PR 16-MAY-2002; 2002US-0381038P.  
XX PR 17-MAY-2002; 2002US-0381042P.  
XX PR 28-MAY-2002; 2002US-0383656P.  
XX PR 29-MAY-2002; 2002US-0383831P.  
XX PR 25-JUN-2002; 2002US-0391335P.  
XX PR 01-OCT-2002; 2002US-00262511.  
XX PA (CURA-) CURAGEN CORP.  
XX PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
XX PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
XX PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
XX PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CBA, Shenoy SG;  
XX PI Shimketa RA, Rothenberg WF, Leach MD, Agee ML, Berghs C, Dippio VA;  
XX PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;  
XX WPI; 2003-381626/36.  
XX P-PSDB; ADA05680.  
XX PT New NOVX polypeptides and nucleic acids, useful for diagnosing,  
XX PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
XX PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
XX PT pharmacogenomics.  
XX PS Claim 20; Page 134; 586pp; English.  
XX PS  
XX CC The present invention describes NOVX proteins, where X can be 1 to 55  
XX CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide  
XX CC described above and a carrier; (2) a kit comprising, in one or more  
XX CC containers, the composition described above; (3) an isolated nucleic acid  
XX CC molecule which encodes a NOVX protein of the invention; (4) a vector  
XX CC comprising the nucleic acid molecule described above; (5) a cell  
XX CC comprising the above vector; (6) an antibody that immunospecifically  
XX CC binds to the polypeptide described above; (7) methods for determining the  
XX CC presence or amount of the above polypeptide or nucleic acid molecule in a  
XX CC sample; (8) methods for determining the presence of or predisposition to  
XX CC a disease associated with altered levels of expression of the above  
XX CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
XX CC method of identifying an agent that binds to the polypeptide described  
XX CC above; (10) a method for identifying a potential therapeutic agent for  
XX CC use in treating a pathology that is related to an aberrant expression or  
XX CC aberrant physiological interactions of the polypeptide; (11) a method of  
XX CC screening for a modulator of activity or of latency or predisposition to  
XX CC a pathology associated with the polypeptide; (12) a method for modulating  
XX CC the activity of the polypeptide described above; (13) methods of treating  
XX CC or preventing a pathology associated with the above polypeptide in a  
XX CC mammal; and (14) a method for producing the above polypeptide. NOVX  
XX CC sequences have antidiabetic, anorectic, antibacterial, virucide,

CC immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian  
CC and antilipaeamic activities, and can be used in gene therapy. The  
CC polypeptide is useful in manufacturing a medicament for treating a  
CC syndrome associated with a human disease. The polypeptide or the nucleic  
CC acid molecule may be used to diagnose, treat or prevent metabolic  
CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
CC disease, immune disorders, haematopoietic disorders and various  
CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
CC probes, in chromosome mapping, tissue typing, preventive medicine and  
CC pharmacogenomics. The present sequence encodes a human NOVX protein from  
CC the present invention.  
XX SQ Sequence 1740 BP; 389 A; 476 C; 530 G; 345 T; 0 U; 0 Other;  
  
Alignment Scores: 3.33e-235 Length: 1740  
Pred. No.: 3025.00 Matches: 562  
Score: 3025.00  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0  
  
US-10-631-958-11 (1-562) x ADA05679 (1-1740)  
  
Qy 1 HisGluAlaAlaAenGlyProAlaProLeuGlyValArgAlaProProAlaTrpArgThr 20  
Db 1 CACGAGGCCGCTAACGCTCCGGCGCCCTCGGGCTCGCGGCCGCCAGCTGGCGGACG 60  
  
Qy 21 SerProAlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrp 40  
Db 61 AGCCCGCGCGGAGATGGGGCGACGCGGGCGCGGAGCGCTGCAATCGTCTGTGG 120  
  
Qy 41 ValysGlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrp 60  
Db 121 GTGAAGCAGCAGCGCTGCGCGCTGAGCTCGAGCCCGCGGGCTCTGCTGCGCTGTGG 180  
  
Qy 61 ArgSerProGlyProGlyAlaGlyAlaProGlyAlaAlaAspAlaCysSerValProValSer 80  
Db 181 CGGAGCCCGGGCGCGGAGCGCGGCCCGCGGTGCTGATGCTGCTGCTGCTGCTATCT 240  
  
Qy 81 GluIleIleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrp 100  
Db 241 GAGATCATCGCGTTGAGGAAACAGACGTTTACAGGAAACATCAAGGACGTCGAAAAATGG 300  
  
Qy 101 GlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgArgHis 120  
Db 301 CAGAAAAATGAAAAAGCCTTACGCTTTTACAGTTTACAGTTTAAAGAGACGACGCGCAC 360  
  
Qy 121 ArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrp 140  
Db 361 CGCTGGGAAGTGGGCGCAGGTGACTTCTGCTGTCAGAGGAGCAGCTGTGTCACTTGTGG 420  
  
Qy 141 LeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuVal 160  
Db 421 CTGCAGACCCCTGCGGAGATGCTGAGAGTGTGAGAGTGTGAGAGTGTGAGTGTGTA 480  
  
Qy 161 PheIleAsnProPheGlyGlyLysGlyGlnClyLysArgIleTyrGluArgLysValAla 180  
Db 481 TTTATCAACCCGTTTGGAGGAAAGGACAGGCAAGCGGATATATGAAGAAAAAGTGTGCA 540  
  
Qy 181 ProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGln 200  
Db 541 CCACGTGTTACCTTAGCCTCCATCACCCTGACATCATCGTTACTGAACATGCTTAATCAG 600  
  
Qy 201 AlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGly 220  
Db 601 GCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGACGATCGTCTGTGTGCGGC 660  
  
Qy 221 GlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAla 240  
Db 661 GGAGATGGTATGTTTACGCGAGGTCTGCACTGCTGATTTGGAGACGACGAGAGCGGCC 720

```
Qy 241 GlyValAspGlnAsnHisProHargAlaValLeuValProSerSerLeuArgIleGlyIle 260
Db 721 GGGGTGCGACAGAAACACCCCGGCTGTGCTGGTCCCGAGTAGCTCCGATGGAATC 780

Qy 261 IleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGlu 280
Db 781 ATTCGCGCAGGTCACCGACTGCTGTGTACTCCAGCTGGGACACAGGACGACAGAA 840

Qy 281 ThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHis 300
Db 841 ACCTCGGCGTGCATATGCTTTGGGACTCGCTGGCCATGGATGTCTCCTCAGTCCAC 900

Qy 301 HisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAsp 320
Db 901 CACAACAGCACACTCCTTCGCTACTCCGCTGCCCTGGCTGGGCTACGGCTTCTACGGGAC 960

Qy 321 IleIleLysAspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGly 340
Db 961 ATCATCAAGGACAGTGAAGAAACGGTGGTGGTCTTGGCAGATACGACTTTTCAGGT 1020

Qy 341 LeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGln 360
Db 1021 TAAAGACCTTCCTCCACCACTGCTATGAAGGACAGTGTCTTCTCCTCGTCACAA 1080

Qy 361 HisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArg 380
Db 1081 CACACGGTGGGATCTCCAAAGGATAGGAAGCCCTCCCGGCGAGGATGCTTTGTTGCAGG 1140

Qy 381 GlnSerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAla 400
Db 1141 CAAACAGCAGCAGCTGGAGGAGGAGGAGAAAGACACTGTATGGTTTGAAGCTCGG 1200

Qy 401 GluAspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsn 420
Db 1201 GAGGACGTGGAGAGTGGCAAGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAAC 1260

Qy 421 MetSerCysAlaCysArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAsp 440
Db 1261 ATGCTCTGTCTGTGTGCGCGGAGCCCGAGGGCCCTCTCCCGGCTGCCCACTTGGGAGAC 1320

Qy 441 GlySerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeu 460
Db 1321 GGGTCTTTGACCTCATCTCATCCGGAATGCTCAGGTTCAATTTTCTGAGATTCTC 1380

Qy 461 IleArgHisThrAsnGlnAspGlnPheAspPheThrPheValGluValTyrArgVal 480
Db 1381 ATCAGGCACACCAACAGCAGGACAGTGTGACTTCACTTTTGTGGAAGTTTATCGCGTC 1440

Qy 481 LysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGly 500
Db 1441 AAGAAATTTCCAGTTTACGTCGAAGCACATGGAGGATGAGGACAGCGACCTCAAGGAGGGG 1500

Qy 501 GlyLysLysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSer 520
Db 1501 GGGAAAGACGCTTTGGGCACATTTGACAGCAGCCACCTCTCTGCTGTCACCGCTCTCC 1560

Qy 521 AsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgVal 540
Db 1561 AACAGCTCTTGAACTGCGAGGGGAGGTCTCTGCACAGCCCTGCCATCGAGTCCAGAGTC 1620

Qy 541 HisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSer 560
Db 1621 CACTGCCAGCTGGTTCGACTCTTTTCAGCAGGAATTTGAAGAGAAATCCGAAGCCAGACTCA 1680

Qy 561 HisSer 562
Db 1681 CACAGC 1686

RESULT 2
ID ABL40828
XX ABL40828 standard; cDNA; 4413 BP.
AC ABL40828;
```

```
XX 03-JUL-2002 (first entry)
DT Human sphingosine kinase-like protein encoding cDNA.
DE
XX Human sphingosine kinase-like protein; intracellular signalling; gene;
KW cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;
KW autoimmune disease; rheumatoid arthritis; Parkinson's disease; ss.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1689
FT /tag= a
FT /product= "sphingosine kinase-like protein"
FT /note= "see ABB07857"
FT CDS 76..1689
FT /tag= b
FT /product= "sphingosine kinase-like protein"
FT /note= "see ABB07856"
XX
XX WO200228906-A2.
XX
XX 11-APR-2002.
XX
XX 05-OCT-2001; 2001WO-EP011516.
PF
XX
XX 06-OCT-2000; 2000US-0238005P.
PR
XX 23-AUG-2001; 2001US-0314113P.
XX
XX (FARB ) BAYER AG.
PA
XX
XX Kossida S, Encinas J;
XX WPI: 2002-340094/37.
XX P-PSDB; ABB07856, ABB07857.
XX
XX New reagent for modulating the activity of sphingosine kinase-like
PT protein polypeptide or polynucleotide and treating cancer, asthma,
PT allergy, an autoimmune disease, or a central or peripheral nervous system
PT disorder.
XX
XX Claim 1; Fig 9; 120pp; English.
XX
XX The invention relates to a human sphingosine kinase-like protein. The
CC polypeptide can be expressed by standard recombinant methodology. The
CC sphingosine kinase-like protein and gene can be used to regulate
CC intracellular signalling and consequently cell proliferation and
CC apoptosis. Such regulation is useful for treating cancer, allergies (e.g.
CC asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and
CC peripheral nervous system disorders (e.g. Parkinson's disease). The
CC present sequence represents the human sphingosine kinase-like protein
CC encoding cDNA
XX
XX Sequence 4413 BP; 1020 A; 1107 C; 1207 G; 1079 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 1.14e-234 Length: 4413
Score: 3025.00 Matches: 562
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 6
```

US-10-631-958-11 (1-562) x ABL40828 (1-4413)

```
Qy 1 HisGluAlaAlaAsnGlyProAlaProLeuGlyValArgAlaProAlaTrpArgThr 20
Db 1 CACGAGGCGCGTAAACGGTCCGCGCCCTCGGCTCCGCGCCGCCAGCTGGCGGACG 60
Qy 21 SerProAlaAlaGluMetGlyAlaThrGlyValAlaGluProLeuGlnSerValLeuTrp 40
Db 61 AGCCCGCGCGGAGATGGGGGCGGACGCGGGCGCGGAGCCGCTGCAATCGTGTGTGG 120
```

```
Qy 41 VallysGlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuLeuArgTyrTrp 60
Db 121 GTGAAGCAGCAGCGCTGCGCGCTGAGCGTGGAGCCCGCGCGGCTCTGCTGCGCTGGTGG 180
Qy 61 ArgSerProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSer 80
Db 181 CGGAGCCCGGGCCCGGAGCCGCGCCCGCGCGCGATGCTGCTCTGCTGCTGTATCT 240
Qy 81 GluIleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrp 100
Db 241 GAGATCATCGCGGTGAGGAACAGACGTTTACCGGAAACATCAAGGACGTGGAAAAATGG 300
Qy 101 GlnLysMetGlnLysProTyrAlaPheThrValHisCysValLysArgAlaArgArgHis 120
Db 301 CAGAAATGGAAAGCCTTAGCTTTTACAGTTTCACTGTGTAAAGAGACGACGAGGCAC 360
Qy 121 ArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGlnGlnLeuCysHisLeuTrp 140
Db 361 CGCTGGAAGTGGCGCAGGTGACTTTCTGGTGTCCAGAGGAGCAGCTGTGTCACTTGTGG 420
Qy 141 LeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuVal 160
Db 421 CTGCAGACCCCTGCGGAGATGCTGGAAAGCTGACGTCAGACCAAGCATTTACTGGTA 480
Qy 161 PheIleAenProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAla 180
Db 481 TTTATCAACCGTTTGAGGNAAGGACNAGCGCAGCGGATATATGAAGAAAGTGGCA 540
Qy 181 ProLeuPheThrLeuAlaSerIleThrThraspileleValThrGluHisAlaAsnGln 200
Db 541 CCACGTTCACCTTAGCCTCCATCACCACATGACATCGTTACTGAACATGCTTAATCAG 600
Qy 201 AlalysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGly 220
Db 601 GCCAAGGAGACTGTGTATGAGATTAAATACAAATAACGACGGCATCGTGTGTGCGC 660
Qy 221 GlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAla 240
Db 661 GGAGATGGTATGTCGCGAGGTGCTGCACGCTCTGATTGGAGGACGACGAGGCGCC 720
Qy 241 GlyValAspGlnAenHisProArgAlaValLeuValProSerSerLeuArgIleGlyIle 260
Db 721 GGGGTGACACGAGAACCCCGCGGCTGTGTGGTCCCGACGTAGCTCCGGAATGGAATC 780
Qy 261 IleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGlu 280
Db 781 ATTCGCGGCTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCTCAGTCCAC 840
Qy 281 ThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHis 300
Db 841 ACCTCGGCGTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCTCAGTCCAC 900
Qy 301 HisAenSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAsp 320
Db 901 CACAACAGCACACTCCTTGGTACTCGCGTGTCCCTGTGGCTACGGCTTCTACGGGGAC 960
Qy 321 IleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGly 340
Db 961 ATCATCAAGGACATGTGAAGAAACGTTGGTGTGGTCTTGCAGATACGACTTTTCAGGT 1020
Qy 341 LeuLysThrPheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAlaGln 360
Db 1021 TTAAGACCTTCTCTCCACCATGTGTATGAAGGACAGTGTCTTCTCTCCCTGCACAA 1080
Qy 361 HisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArg 380
Db 1081 CACACGTTGGATCTCCAAGGATAGGAAGCCCTGCCGGGAGGATGCTTTGTTGTCAGG 1140
Qy 381 GlnSerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAla 400
Db 1141 CAAAGCAAGCAGCAGCTGGAGGAGGACGAAGAAAGCAGCTGTATGTTTGAAGCTGCG 1200
```

```
Qy 401 GluAspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsn 420
Db 1201 GAGCAGTGGAGAGTGGCAAGTCGTGTGTGGGAAGTTTCTGGCCATCAATGCCACAAAC 1260
Qy 421 MetSerCysAlaCysArgArgSerProArgGlyLysSerProAlaAlaHisLeuGlyAsp 440
Db 1261 ATCTCTCTGTCTTGTGCGCGAGCCCGAGGGCTCTCCCGGCTGCCCATTTGGGAGAC 1320
Qy 441 GlySerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeu 460
Db 1321 GGGTCTTCTGACCTCATCTCATCCGGAATGTCCAGGTTCATTTTCTGAGATTTCTC 1380
Qy 461 IleArgHisThrAenGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgVal 480
Db 1381 ATCAGGCACACCAACCCAGCAGGACCAAGTTTGACTTCACTTTTTGTGAAGTTTATCGCGTC 1440
Qy 481 LysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGly 500
Db 1441 AAGAAATTCAGTTTACGTGCAAGCACATGAGAGGATGAGGACGCGACCTCAAGGAGGG 1500
Qy 501 GlyLysLysArgPheGlyHisIleCysSerSerHisProSerCysCysCysThrValSer 520
Db 1501 GGGAAAGAGCGCTTGGGCACATTTGCAGCAGCACCCCTCTCTGCTGCTCACGCTCC 1560
Qy 521 AsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgVal 540
Db 1561 AACAGCTCCTGGAACCTGCGACGGGAGGTCTCTGCACAGCCCTGCCATCGAGGTCA 1620
Qy 541 HisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAenProLysProAspSer 560
Db 1621 CACTGCCAGCTGTGTTGCACTCTTTTGACGAGGAATTAAGAGAATTCGAAGCCAGACTCA 1680
Qy 561 HisSer 562
Db 1681 CACAGC 1686
RESULT 3
ADJ96598
ID ADJ96598 standard; DNA; 4429 BP.
XX
AC ADJ96598;
XX
XX 06-MAY-2004 (first entry)
XX
DE Human lipid kinase KIAA1646 DNA SeqID 55.
XX
KW gene; ds; kinase; human; SNP; single nucleotide polymorphism;
KW tyrosine protein kinase; serine/threonine protein kinase; ERK; STK;
KW gene therapy; cancer; immune-related disease; cardiovascular disease;
KW brain; neuronal associated disease; metabolic; inflammatory disorder;
KW cytosolic; neuroprotective; immunomodulator; antiinflammatory;
KW lipid kinase; KIAA1646.
XX
OS Homo sapiens.
OS 39.
XX
PH Key Location/Qualifiers
FT variation /*tag= a
FT replace(2391,g)
FT /*standard_name= "Single nucleotide polymorphism"
FT variation /*tag= b
FT replace(2577,g)
FT /*standard_name= "Single nucleotide polymorphism"
FT variation /*tag= c
FT replace(3020,c)
FT /*standard_name= "Single nucleotide polymorphism"
FT variation /*tag= d
FT replace(3769,c)
FT /*standard_name= "Single nucleotide polymorphism"
FT variation /*tag= e
FT replace(4272,g)
FT /*standard_name= "Single nucleotide polymorphism"
XX
```

PN WO2004006838-A2.  
 XX 22-JAN-2004.  
 XX 15-JUL-2003; 2003WO-US021730.  
 XX 15-JUL-2002; 2002US-0395632P.  
 XX (SUGEN-) SUGEN INC.  
 XX Whyte D, Manning G, Caenepeel S;  
 XX WPI: 2004-122753/12.  
 DR P-PSDB; ADJ96664.  
 XX  
 PT New nucleic acid molecule encoding a kinase polypeptide, useful for  
 PT preparing a composition for treating diseases or disorders, e.g., cancer,  
 PT or neurological, immunological or inflammatory disorders.  
 XX  
 PS Example 1; SEQ ID NO 55; 366pp; English.  
 XX  
 CC This invention relates to a novel isolated, enriched or purified nucleic  
 CC acid molecule that encodes a kinase polypeptide. Specifically, it relates  
 CC to human tyrosine and serine/threonine protein kinases (PTK's and STK's),  
 CC as well as protein kinase-like enzymes. The present invention describes  
 CC screening methods to identify agonists, antagonists and antibodies that  
 CC can be used to modulate the activity or function of the mammalian kinase  
 CC enzymes. As such, these compositions can be used for gene therapy  
 CC purposes to treat diseases or disorders including cancer, immune-related  
 CC diseases, cardiovascular disease, brain or neuronal associated disease,  
 CC metabolic and inflammatory disorders. Accordingly, they exhibit  
 CC cytosolic, neuroprotective, immunomodulator and anti-inflammatory  
 CC activities. This polynucleotide sequence is a human kinase DNA sequence  
 CC of the invention.  
 XX  
 SQ Sequence 4429 BP; 1020 A; 1115 C; 1213 G; 1081 T; 0 U; 0 Other;  
 XX  
 Alignment Scores:  
 Pred. No.: 1.3e-233 Length: 4429  
 Score: 3012.00 Matches: 560  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 99.5% Indels: 0  
 DB: 12 Gaps: 0  
 US-10-631-958-11 (1-562) x ADJ96598 (1-4429)  
 QY 3 AlaAlaAsnGlyProAlaProLeuGlyValArgAlaProProAlaTrpArgThrSerPro 22  
 DB 23 GCCGTAACGGTCCGGCGCCCTCGGGCGTCCGGCGCCCGCCAGCCTGGCGAGAGCCCG 82  
 QY 23 AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpVallys 42  
 DB 83 GCGGCGGAGATGGGGCGACGGGGCGGGCGGAGCGCTGCATCCGTCTGTGGGTGAAG 142  
 QY 43 GlnGlnArgCysAlaValSerLeuGluProAlaAlaGluLeuLeuArgTrpArgSer 62  
 DB 143 CAGCAGCGTGGCGCGTGGAGCTGGAGCCCGCGCGGCTCTCTCGCTGGTGGCGGAGC 202  
 QY 63 ProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluile 82  
 DB 203 CCGGGGCGCGAGCGCGCCCGCGCGGATGCTCTGTCTGCTGCTGATCTCAGATC 262  
 QY 83 IleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLys 102  
 DB 263 ATCCGCGTTGAGAAACAGACGTTTCACGGGAAACATCAAGGCGATGGGAAATGGCAGAA 322  
 QY 103 MetGluLysProTyAlaPheThrValHisCysValLysArgAlaArgHisArgTrp 122  
 DB 323 ATGGAAAGCCCTTACGCTTTTACAGTTTCACTGTGTAAAGAGAGACGACGCGCAGCTGG 382  
 QY 123 LysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuLeuCysHisLeuTrpLeuGln 142

Db 383 AAGTGGGCGAGGTGACTTTCTGTGTCCAGAGAGAGAGCTGTGTCACTTGTGGCTGCAG 442  
 QY ThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIle 162  
 Db 443 ACCCTGGGGAGATGCTGGAGAAGCTGACGTCCAGACCAGCAATTTACTGGTATTATC 502  
 QY AsnProPheGlyGlyLysGlyGlnGlyArgIleTyrGluArgLysValAlaProLeu 182  
 Db 503 AACCCGTTTGGAGGAAAGGACCAAGGCAAGCGGATATATGAAGAAAAAGTGGCCACTG 562  
 QY PheThrLeuAlaSerIleThrAspIleValThrGluHisAlaAsnGlnAlaLys 202  
 Db 563 TTCACCTTAGCTCCATCACCACCTGACATCATCTGTACTGAACATGCTAATCAGGCCAAG 622  
 QY GluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyAsp 222  
 Db 623 GAGACTCTGTATGAGATTAAATAGACAAATACGACGCGCATCGTGTGTGGCGGAGAT 682  
 QY GlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyVal 242  
 Db 683 GGTATGTTTCAGCGAGGTGCTGCAGCGTCTGATTGGGAGGACGAGAGCGCGGGTTC 742  
 QY AspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIlePro 262  
 Db 743 GACCAAGAACACCCCGGGCTGTGTGGTCCCGAGTAGCTCCGGATTGGAAATCATTC 802  
 QY AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer 282  
 Db 803 GCAGGGTCAACGGACTGCGTGTGTGTACTCCACCGTGGGCGACGACGACGAGAACCTCG 862  
 QY AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsn 302  
 Db 863 CGCTGTCATATCGTTGTTGGGAGCTCGCTGGCCATGGATGTGTCTCAGTCCACCAAC 922  
 QY SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIle 322  
 Db 923 AGCACACTCTTCGCTACTCTCGTGTCTGTGGCTACGGCTTCTACGGGAGCATATC 982  
 QY LysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys 342  
 Db 983 AAGGACAGTGAAGAAGAACGGTGGTGTGGTCTTGCAGATACGACTTTTTCAGGTTTAAAG 1042  
 QY ThrPheLeuSerHisHisCysTyrGlyThrValSerPheLeuProAlaGlnHisThr 362  
 Db 1043 ACCTTCTCTCCACCACTGCTATGAAGGGACAGTGTCTTCTCTCTCCCTGCAACACACG 1102  
 QY ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer 382  
 Db 1103 GTGGATCTCCAGGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTGGAGGCAAGC 1162  
 QY LysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp 402  
 Db 1163 AAGCAGCAGCTGGAGGAGGAGCAGAGAAAGACACTGTATGTTTGGAAAGCTGCGGAGAC 1222  
 QY ValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSer 422  
 Db 1223 GTGGAGGAGTGGCAAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCC 1282  
 QY CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer 442  
 Db 1283 TGTGCTTGTCCCGAGAGCCCGAGGGCTCTCCCGGGCTGCCACTTGGGAGAGCGGGTCT 1342  
 QY SerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArg 462  
 Db 1343 TCTGACCTCATCTCATCCGGAATGCTCCAGGTTCATATTTCTGAGATTTCTCATCAGG 1402  
 QY HisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLys 482  
 Db 1403 CACACCAACGACGAGGACGAGTTTCACTTCACTTTTGTGAAGTTTATTCGGTGTCAAGAAA 1462  
 QY PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLys 502  
 Db 1463 TTCCAGTTTACGTGGAAGCACATGAGGATGAGGACGACGACCTCAAGAGGGGGGGAAG 1522

Qy 503 LysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSer 522  
 Db 1523 AAGCGCTTTGGGACATTTGCAGCAGCACCCTCTGCTGCTGACCGTCTCCACAGC 1582  
 Qy 523 SerTrpIenCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCys 542  
 Db 1583 TCCTGGAACTCGACGGGGAGGTCTGCACAGCCCTGCCATCGAGGTTCAGAGTCCACTGC 1642  
 Qy 543 GlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProIysProAspSerHisSer 562  
 Db 1643 CAGCTGGTTCAGTCTCTTGCACGAGGAATTGAAGAGAATCCGAAGCCAGACTCACAGC 1702

## RESULT 4

ADP55247

ID ADP55247 standard; cDNA; 4445 BP.

XX AC ADP55247;

XX DT 18-NOV-2004 (first entry)

XX DE Human PRO cDNA sequence SEQ ID NO:1223.

XX KW human; PRO; immune related disease; inflammatory immune response;  
 KW immune response stimulation; antiallergic; antianaemic; antiarthritic;  
 KW antiasthmatic; antidiabetic; antinflammatory; antipsoriatic;  
 KW antirheumatic; antichyroid; CNS; dermatological; gastrointestinal;  
 KW haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;  
 KW nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;  
 KW virucide; gene therapy; gene; ss.

XX OS Homo sapiens.

XX PN WO2004039956-A2.

XX PD 13-MAY-2004.

XX PF 28-OCT-2003; 2003WO-US034381.

XX PR 29-OCT-2002; 2002US-0422472P.

XX PA (GSTH ) GENENTECH INC.

XX PI Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;

XX PI Wood WJ, Wu TD;

XX DR WPI; 2004-376182/35.

XX DR P-PSDB; ADP55248.

XX PT New PRO polynucleotides and polypeptides, useful in useful in diagnosing  
 PT and treating an immune related disease, e.g. systemic lupus  
 PT erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in  
 PT stimulating an immune response.

XX PS Claim 2; SEQ ID NO 1223; 3009pp; English.

XX PS The present invention describes an isolated PRO nucleic acid (1). Also  
 CC described: (1) a vector comprising (1); (2) a host cell comprising the  
 CC vector of (1); (3) a process for producing a PRO polypeptides; (4) an  
 CC isolated PRO polypeptide; (5) a chimeric molecule comprising the  
 CC polypeptide of (4) fused to a heterologous amino acid sequence; (6) an  
 CC antibody which specifically binds to a polypeptide of (4); (7) a  
 CC composition of matter comprising a polypeptide of (4), an agonist or  
 CC antagonist of the polypeptide or an antibody that binds to the  
 CC polypeptide in combination with a carrier; (8) an article of manufacture  
 CC comprising a container, a label on the container and a composition of  
 CC matter of (7); (9) a method of treating an immune related disease in a  
 CC mammal; (10) a method for determining the presence of a PRO polypeptide  
 CC in a sample suspected of having the polypeptide; (11) a method of  
 CC diagnosing an immune related disease or an inflammatory immune response  
 CC in mammal; (12) a method of identifying a compound that inhibits or  
 CC mimics the activity of or expression of a gene encoding a PRO polypeptide  
 CC ; and (13) a method of stimulating the immune response in a mammal. The

CC PRO sequences have antiallergic, antianaemic, antiarthritic,  
 CC antiasthmatic, antidiabetic, antinflammatory, antipsoriatic,  
 CC antirheumatic, antichyroid, CNS, dermatological, gastrointestinal,  
 CC haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,  
 CC nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and  
 CC virucide activities, and can be used in gene therapy. The nucleic acid  
 CC (1) and the encoded polypeptides, compositions, kits and methods are  
 CC useful in diagnosing and treating an immune related disease and in  
 CC stimulating an immune response. The present sequence represents a human  
 CC PRO nucleotide sequence from the present invention.

XX SQ Sequence 4445 BP; 1018 A; 1134 C; 1216 G; 1077 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 1-31e-233 Length: 4445  
 Score: 3012.00 Matches: 560  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 99.57% Indels: 0  
 DB: 13 Gaps: 0

US-10-631-958-11 (1-562) x ADP55247 (1-4445)

Qy 3 AlaAlaAsnGlyProAlaProLeuGlyValArgAlaProProAlaTrpArgThrSerPro 22

Db 55 GCGCGCTAACGCTCCGCGCGCCCTCGCGCTCCGCGCGCCCGCCAGCTGGCGGAGCGCG 114

Qy 23 AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLys 42

Db 115 GCGGCGAGATGGGGGCGAGCGGGGGCGGCGCGCGCTGCATCCGTCGTGGGTGAAG 174

Qy 43 GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSer 62

Db 175 CAGCAGCGCTGCGCGCTGAGCGCTGGAGCCGCGCGGGCTCTGCTGGCTGGCGGAGC 234

Qy 63 ProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIle 82

Db 235 CCGGGCGCGGAGCGCGCGCGCGCGCGGATGCCCTGCTCTGCTGCTGTATCTGAGATC 294

Qy 83 IleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLys 102

Db 295 ATGCCGTTGAGGAACACAGACGTTTCACGGGAACATCAAGGCAGTGGAAATGGCAGAAA 354

Qy 103 MetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTyr 122

Db 355 ATGGAAGAGCGCTTACGCTTTTACAGTTTCACTGTGTAAAGAGAGCAGCAGCGCCTGG 414

Qy 123 LysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGln 142

Db 415 AAGTGGCGCGAGGTGACTTTCTGGTGTCCAGAGGAGCAGCTGTCTCACTTGTGGCTGCAG 474

Qy 143 ThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIle 162

Db 475 ACCCTGCGGAGATGCTGGAGAAGCTGACCTCCAGACCAAGCATTTTACTGGTATTATC 534

Qy 163 AsnProPheGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeu 182

Db 535 AACCCGTTTGGAGAAAAGGACCAAGCAAGCGGATATATCAAGAAAGTGGCACCCTG 594

Qy 183 PheThrLeuAlaSerIleThrThrAspIleValThrGluHisAlaAsnGlnAlaLys 202

Db 595 TTCACCTTAGCCCTCCATCACCACTCATCATCTTACTGAAACATGCTAATACAGGCCAAG 654

Qy 203 GluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAsp 222

Db 655 GAGACTCTGATGAGATTAACATAGACAATATACAGCGCATGCTGTGTGCGCGGAGAT 714

Qy 223 GlyMetPheSerGluValLeuHisGlyIleGlyArgThrGlnArgSerAlaGlyVal 242

Db 715 GGTATGTTACGCGAGGTGCTGCGCGTCTCATTTGGGAGGACGACGAGCGCGCGGGTTC 774

Qy 243 AspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIlePro 262

Db 775 GACGAGAACCCCGGGCTGTGCTGGTCCAGTAGCTCCGATTCGGATCATTCCTCC 834  
Qy 263 AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer 282  
Db 835 GCAGGGTCAACGAGCTGGCTGTGTACTCCACCGTGGGCACGACGACGAGAACTTCG 894  
Qy 283 AlaLeuHisIleValValGlyValSerLeuAlaMetAspValSerSerValHisIleAsn 302  
Db 895 GCGCTGCATATCGTTTGTGGGACATCGCTGGCCATGGATGTCTCTAGTCCACCAAC 954  
Qy 303 SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIle 322  
Db 955 AGCACATCTCTCGTACTCGTGTCTCTGCTGGGCTACGGCTTACGGGGACATCATC 1014  
Qy 323 LysAspSerGluLysIleValGlyTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys 342  
Db 1015 AAGGACAGTGAAGAAGACGGTGGTGTGGTCTTGGCAGATACGATTTTCAGGTTAAAG 1074  
Qy 343 ThrPheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr 362  
Db 1075 ACCTTCTCTCCACACATGCTATGAAGGACAGTGTCTCTCTCCCTGCACACACACG 1134  
Qy 363 ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer 382  
Db 1135 GTGGGATCTCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTGTGTGAGGCAAGC 1194  
Qy 383 LysGlnGlnLeuGluGluGlnLysIleValLeuTyrGlyLeuGluAlaGluAsp 402  
Db 1195 AAGCAGCAGCTGGAGGAGAGCAGAAAGCACTGTATGTTTGAAGCTGCGGAGGAC 1254  
Qy 403 ValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSer 422  
Db 1255 GTGGAGGAGTGGCAAGTGTCTGTGGGAGTTTCTGGCCATCAATGCCACAAACATGTCC 1314  
Qy 423 CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer 442  
Db 1315 TGTGCTTGTGCGCGGAGCCAGGGGCTCTCCCGGGTGGCCACTTGGAGACGGGTCT 1374  
Qy 443 SerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArg 462  
Db 1375 TCTGACCTCATCTCATCGGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGG 1434  
Qy 463 HisThrAsnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLys 482  
Db 1435 CACACCAACACGACGACGAGTTTCACTTCTTGTGAGTTTATCGCGTCAAGAAA 1494  
Qy 483 PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLys 502  
Db 1495 TTCCAGTTTACGTCGACGACATGAGGATGAGGACGAGGACCTCAAGAGGGGGGGAAG 1554  
Qy 503 LysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSer 522  
Db 1555 AAGCGCTTTGGGCACATTTGCAGCAGCAGCCCTCTCTGCTGTCGACCGTCTCCACAGC 1614  
Qy 523 SerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCys 542  
Db 1615 TCCTGGAATCGGATGGGAGGTCCTGCACAGCCCTGCCATCGAGGTGAGATCCACTGC 1674  
Qy 543 GlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 562  
Db 1675 CAGTGGTTCGACTCTTTGACAGGAGATTTGAAGAGATTCGACCCAGATCCACAGC 1734

## RESULT 5

ABA96945

ID ABA96945 standard; cDNA; 4463 BP.

XX

AC ABA96945;

XX

DT 20-MAY-2002 (first entry)

XX

DE Human ceramide kinase hCERK1-encoding cDNA.

XX

KW Human; ceramide kinase; hCERK1; drug screening; gene therapy;

KW neurological disease; inflammation; human immunodeficiency virus;  
KW HIV infection; type 2 diabetes; obesity; sepsis; arteriosclerosis;  
KW cancer; neuroprotective; anti-inflammatory; anti-HIV; antidiabetic;  
KW anorectic; antibacterial; antiseptic; antiarteriosclerotic; cytostatic;  
KW gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FH CDS 124..1737

FT /\*tag= a

XX /product= "Human ceramide kinase hCERK1"

FN WO200196575-A1.

XX 20-DEC-2001.

XX 11-JUN-2001; 2001WO-JP004889.

XX 14-JUN-2000; 2000JP-00178039.

XX (SANY ) SANKYO CO LTD.

XX Sugiuira M, Kono K, Kohama T;

XX WPI; 2002-179513/23.

DR P-PSDB; AAM49115.

XX Human ceramide kinase gene and the enzyme encoded by it for screening  
XX substances as drugs for neurological, inflammatory and other disorders.

PS Claim 5; Page 46-53; 61pp; Japanese.

XX This sequence represents cDNA encoding a human ceramide kinase designated  
XX hCERK1. The invention relates to hCERK1, nucleic acids encoding it,  
XX expression vectors and host cells containing hCERK1 nucleic acids, the  
XX recombinant production of hCERK1 and antibodies specific for hCERK1. The  
XX invention also encompasses methods of isolating hCERK1 from samples, the  
XX use of hCERK1 in drug screening, and the use of hCERK1 nucleic acid  
XX sequences in gene therapy. hCERK1 mediates the ATP-dependent 1-  
XX phosphorylation of ceramides and can be used to screen for therapeutic  
XX and preventive agents for a wide range of disorders. Such disorders  
XX include neurological disease, inflammation, human immunodeficiency virus  
XX (HIV) infection, type 2 diabetes, obesity, sepsis, arteriosclerosis and  
XX cancer

SQ Sequence 4463 BP; 1026 A; 1135 C; 1217 G; 1084 T; 0 U; 1 Other;

## Alignment Scores:

Pred. No.: 5.84e-233 Length: 4463  
Score: 3004.00 Matches: 558  
Percent Similarity: 99.82% Conservative: 1  
Best Local Similarity: 99.64% Mismatches: 1  
Query Match: 99.31% Indels: 0  
DB: 6 Gaps: 0

US-10-631-958-11 (1-562) x ABA96945 (1-4463)

Qy 3 AlaAlaGlnGlyProAlaProLeuGlyValArgAlaProProAlaTrpArgThrSerPro 22  
Db 55 GCCGCTAACGGTCGGGGCCCTCGGGTCCGGCGGCCCGCCAGCTGGCGACGAGCCCG 114

Qy 23 AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLys 42  
Db 115 GCGGCGGAGATGGGGCGACGGGGCGCGGAGCCGCTGCAATCCGTGCTGTGGGTGAAG 174

Qy 43 GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSer 62  
Db 175 CAGCAGCGCTGCGCCGTGAGCTGGAGCCCGCGGGCTCTGCTGCGCTGTGGCGGAGC 234

Qy 63 ProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIle 82  
Db 235 CCGGGGCGCGAGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 294



Qy 83 IleAlaValGluGluThrAspValHisGlyIysHisGlnGlySerGlyIysTrpGlnLys 102  
Db 295 ATCCCGCTTGAGGAAACAGACGTTACGGGAAACATCAGGCACTGGGAAATGCGAGAA 354  
Qy 103 MetGluLysProTyrAlaPheThrValHisCysValLysArgAlaLaArgHisArgTrp 122  
Db 355 ATGGAAGAGCCCTTACGGCTTTTACAGTTCACTGTGTAAAGAGAGCAGCAGCGCCGCTGG 414  
Qy 123 LysTrpAlaGlnValThrPheTrpCysProGluGlnLeuGlnCysHisLeuTrpLeuGln 142  
Db 415 AAGTGGGCGCAGGTGACTTCTGTGTCCAGAGAGCAGCTGTGTCACTTGTGTGCTGCAG 474  
Qy 143 ThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIle 162  
Db 475 ACCCTGGCGGAGATGCTGGAGAGCTGACGTCACAGCCAAAGCAATTTACTGGTATTATC 534  
Qy 163 AsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeu 182  
Db 535 AACCCGTTTGGAGGAAAGGACAAAGCAAGCGGATATATGAAGAAAGTGGCACCACTG 594  
Qy 183 PheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLys 202  
Db 595 TTCACCTTAGCCCTCCATCACCACATGATCATCGTTACTGAACATGCTAATTCAGGCCAAG 654  
Qy 203 GluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAsp 222  
Db 655 GAGACTCTGTATGAGATTAAATAGACAAATACACGGCATCGTCTGTGTCGGCGGAGAT 714  
Qy 223 GlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyVal 242  
Db 715 GGTATGTTACGAGGTTGCTGCACGGTCTGATTTGGGAGCGCAGAGAGCGCGGGTTC 774  
Qy 243 AspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleLeuPro 262  
Db 775 GACCAGAAACACCCCGGGCTGTGCTGCTCCAGTAGTACCTCCGGATTGGGAATCATTTCCC 834  
Qy 263 AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer 282  
Db 835 GCAGGTCACAGGACTCGGTGTGTACTCCACCGTGGGCACGAGCGCAGAAACCTCG 894  
Qy 283 AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsn 302  
Db 895 CGCGCTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTGCTCAGTCCACCAAC 954  
Qy 303 SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIle 322  
Db 955 AGCACCTCTCTCGTACTCGTGTCCGTGCTGGGCTACGGCTTCTACGGGGAATCATC 1014  
Qy 323 LysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys 342  
Db 1015 AAGGACAGTGAGAGAAACGGTGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAG 1074  
Qy 343 ThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr 362  
Db 1075 ACCTTCTCTCCCACTATGATGAAGGACAGTGTCTTCTCCCTCGCAGCAACACACG 1134  
Qy 363 ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer 382  
Db 1135 GTGGGATCTCAAGGATAGGAAGCCCTGCCGGGAGGATGCTTGTGTGGAGGCAAGC 1194  
Qy 383 LysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp 402  
Db 1195 AAGCAGCAGCTGGAGGAGGACAGAGAAACCACTGTATGTTTGGAGCTGCGGAGGAC 1254  
Qy 403 ValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSer 422  
Db 1255 GTGGAGGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCC 1314  
Qy 423 CysAlaCysArgArgSerProArgGluLeuSerProAlaAlaHisLeuGlyAspGlySer 442  
Db 1315 TGTGCTGTCCCGGAGCCCGAGGGGCTCTCTCCCGGCTGCCCATCTTGGGAGAGCGGTCT 1374

Qy 443 SerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArg 462  
Db 1375 TCTGACCTCATCTCTATCCGAAATGCTCCAAAGTTCAATTTTCTGAGATTTCTCATCAG 1434  
Qy 463 HisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLys 482  
Db 1435 CACACCAACAGCAGGACCACTTTCACCTTTCACCTTTGTTGAAGTTTATTCGCGTCAAGAA 1494  
Qy 483 PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyGlyLys 502  
Db 1495 TTCAGTTTACGTCGAAGCACATGGAGATCAGGACAGCCCTCAAGAGGGGGGAG 1554  
Qy 503 LysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSer 522  
Db 1555 AAGCGCTTGGGCACATTTTGCAGCAGCCACCTCTCTGCTGCTGACCGTCTCCACAGC 1614  
Qy 523 SerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCys 542  
Db 1615 TCCTGGAACTGCGATGGGAGGTCTTCGACAGCCCTGCCATCGAGGTCCAGTCCACTGC 1674  
Qy 543 GlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 562  
Db 1675 CAGCTGGTTGCATCTCTTGCACGAGGANTTGAAGAGAAATCCGAAGCCAGACTCACAGC 1734

RESULT 6  
ADN62844  
ID ADN62844 standard; DNA; 1740 BP.  
XX  
AC ADN62844;  
XX  
DT 01-JUL-2004 (first entry)  
XX  
DE Human NOV9a DNA.  
XX  
KW ds; gene; human; NOVX; metabolic disorder; diabetes; obesity;  
KW infectious disease; anorexia; cancer; cancer-associated cachexia;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW immune disorder; haematopoietic disorder; dyslipidaemia;  
KW metabolic syndrome X; wasting disorder.  
XX  
OS Homo sapiens.  
XX  
PN US2004038223-A1.  
XX  
PD 26-FEB-2004.  
XX  
PF 01-OCT-2002; 2002US-00262511.  
XX  
PR 02-OCT-2001; 2001US-0326483P.  
PR 05-OCT-2001; 2001US-0327435P.  
PR 05-OCT-2001; 2001US-0327449P.  
PR 09-OCT-2001; 2001US-0327917P.  
PR 09-OCT-2001; 2001US-0328029P.  
PR 09-OCT-2001; 2001US-0328044P.  
PR 09-OCT-2001; 2001US-0328056P.  
PR 12-OCT-2001; 2001US-0328849P.  
PR 15-OCT-2001; 2001US-0329414P.  
PR 17-OCT-2001; 2001US-0330142P.  
PR 18-OCT-2001; 2001US-0330309P.  
PR 22-OCT-2001; 2001US-0341058P.  
PR 24-OCT-2001; 2001US-0339266P.  
PR 24-OCT-2001; 2001US-0343629P.  
PR 29-OCT-2001; 2001US-0349575P.  
PR 01-NOV-2001; 2001US-0346357P.  
PR 17-APR-2002; 2002US-0373260P.  
PR 19-APR-2002; 2002US-0373815P.  
PR 19-APR-2002; 2002US-0373817P.  
PR 19-APR-2002; 2002US-0373826P.  
PR 19-APR-2002; 2002US-0373884P.  
PR 22-APR-2002; 2002US-0374977P.  
PR 16-MAY-2002; 2002US-0381037P.  
PR 16-MAY-2002; 2002US-0381038P.  
PR 16-MAY-2002; 2002US-0381042P.



PR	17-MAY-2002; 2002US-0381642P.	CC	diagnostic agents for detecting the presence of NOVX in samples. NOVX
PR	28-MAY-2002; 2002US-0383656P.	CC	polypeptides and polynucleotides may be used in this way to prevent,
PR	29-MAY-2002; 2002US-0383831P.	CC	diagnose and treat: metabolic disorders, diabetes, obesity, infectious
PR	25-JUN-2002; 2002US-0391335P.	CC	disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative
XX	(SMIT//) SMITHSON G.	CC	disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,
PA	(MILL//) MILLET I.	CC	haematopoietic disorders, and the various dyslipidaemias, metabolic
PA	(PEYM//) PEYMAN J A.	CC	disturbances associated with obesity, the metabolic syndrome X and
PA	(KEKU//) KEKUDA R.	CC	wasting disorders associated with chronic diseases and various cancers.
PA	(JUQU//) JU J.	CC	They may also be used as antibacterial agents. The present sequence
PA	(LILL//) LI L.	CC	represents DNA encoding a human NOVX protein.
PA	(GUOX//) GUO X.	XX	
PA	(PATT//) PATTURAJAN M.	SQ	Sequence 1740 BP; 388 A; 477 C; 530 G; 345 T; 0 U; 0 Other;
PA	(SPYT//) SPYTEK K A.		
PA	(EDIN//) EDINGER S R.	Alignment Scores:	
PA	(ELLE//) ELLERMAN K.	Pred. NO.: 1.47e-231	Length: 1740
PA	(MALY//) MALYANKAR U M.	Score: 2980.00	Matches: 560
PA	(ORTT//) ORT T.	Percent Similarity: 99.64%	Conservative: 1
PA	(GORM//) GORMAN L.	Best Local Similarity: 99.47%	Mismatches: 1
PA	(ZERH//) ZERHUSEN B D.	Query Match: 98.51%	Indels: 2
PA	(ANDE//) ANDERSON D W.	DB: 12	Gaps: 0
PA	(ZHON//) ZHONG M.	US-10-631-958-11 (1-562) x ADN62844 (1-1740)	
PA	(CATT//) CATTERTON E.		
PA	(JIW//) JI W.	Qy 1 HisGluAlaAlaAsnGlyProAlaProLeuGlyValArgAlaProProAlaTrpArgThr 20	
PA	(MILL//) MILLER C E.	Db 1 CACGAGCCCGCTAACGGTCCGGCGCCCTCGGGCGTCCGGCGCCCGCCAGCTGGCGGACG 60	
PA	(RAST//) RASTELLI L.	Qy 21 SerProAlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrp 40	
PA	(STON//) STONE D J.	Db 61 AGCCCGCGCGGAGATGGGGCGGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120	
PA	(PENA//) PENA C E A.	Qy 41 ValysGlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrp 60	
PA	(SHEN//) SHENOY S G.	Db 121 GTGAAGCAGCAGCGCTGCGCGTGCAGCTGAGCGCCGCGGCGGCTCTGCTGCGCTGGTGG 180	
PA	(SHIM//) SHIMKETS R A.	Qy 61 ArgSerProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSer 80	
PA	(ROTH//) ROTHENBERG M E.	Db 181 CGGAGCCCGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240	
PA	(LEAC//) LEACH M D.	Qy 81 GluIleIleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrp 100	
PA	(AGEE//) AGEE M L.	Db 241 GAGATCATCGCGTGTGAGGAAACAGAGCTTACCGGAAACATCAAGGCGAGTGGAAATGG 300	
PA	(BERG//) BERGHS C.	Qy 101 GlnLysMetGluLysProTyrrAlaPheThrValHisCysValLysArgAlaArgHis 120	
PA	(DIPI//) DIPIPO V A.	Db 301 CAGAAAATGAAAAGCCTTACGCTTTTACAGTTTACAGTTTACAGTTTACAGTTTAC 360	
PA	(EISE//) EISEN A.	Qy 121 ArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuLeuLeuTrp 140	
PA	(GANG//) GANGOLLI E A.	Db 361 CGCTGGGAAGTGGCGCGCAGGTGACTTTCTGGTGTCCAGAGGAGCAGCTGTGTCACTTGTGG 420	
PA	(RIEG//) RIEGER D K.	Qy 141 LeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuVal 160	
PA	(SPAD//) SPADERNA S K.	Db 421 CTGCAGACCCCTGCGGAGATGCTGGAGAGAGCTGAGAGAGCTGAGAGAGAGAGAGAG 480	
XX	Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;	Qy 161 PheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrrGluArgLysValAla 180	
PI	Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malvankar UM;	Db 481 TTTTATCAACCCGTTTGGAGGAAAAGGCAAGGCAAGCGGATATATGAAAGAAAAGTGGCA 540	
PI	Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;	Qy 181 ProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGln 200	
PI	Ji W, Miller CB, Rastelli L, Stone DJ, Pena CE, Shenoy SG;	Db 541 CCACCTGTTTACCTTTAGCCTCCATCACCACCTGACATCATCGTTTACTGAACTGCTAATCAG 600	
PI	Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;	Qy 201 AlaLysGluThrLeuTyrrGluIleAsnIleAspLysTyrrAspGlyIleValCysValGly 220	
PI	Eisen A, Gangolli EA, Rieger DK, Spaderna SK;	Db 601 GCCAAGGAGACTCTGTATGAGATTAAATAGCAAAATACAGCGGCAT-GTCTGTGTGCGC 659	
XX	WPI: 2004-213931/20.	Qy 221 GlyAsp-GlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAl 240	
DR	P-FSDB; ADN62845.	Db 660 GGAGATCGGTATGTTTACGAGGAGGTGCTGCACGGTCTGATTGGGAGGAGCGCAGAGAGCGC 719	
DR	Isolated NOVX polypeptides and nucleic acids, useful for preventing,	Qy 240 aGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIle 260	
XX	diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.		
PT	Claim 20; SEQ ID NO 39; 395pp; English.		
PT	The invention relates to isolated NOVX polypeptides and polynucleotides.		
XX	NOVX polypeptides and polynucleotides are used to prevent, diagnose or		
CC	treat a medical condition in human related to the aberrant expression and		
CC	activity of NOVX polypeptides. For example, NOVX polypeptides and		
CC	polynucleotides may be used to treat disorders associated with decreased		
CC	expression or activity of NOVX by supplementing the patient our		
CC	production or to rectify mutations. Conversely, antisense NA molecules		
CC	may be administered to down regulate expression of NOVX polypeptides by		
CC	binding with the cells own genes and preventing their expression. NOVX		
CC	polynucleotides and complementary sequences may also be used as DNA		
CC	probes in diagnostic assays to detect and quantitate the presence of		
CC	similar sequences in samples, and so which patients may be in need of		
CC	restorative therapy. NOVX polypeptides may also be used as antigens in		
CC	the production of antibodies and in assays to identify modulators		
CC	(agonists and antagonists) of the expression and activity of NOVX. The		
CC	anti-NOVX polypeptide antibodies, agonists and antagonists may also be		
CC	used to modulate NOVX polynucleotide expression and activity of NOVX		
CC	polypeptides. The anti-NOVX polypeptide antibodies may also be used as		



DB:	8	Gaps:	1
US-10-631-958-11	(1-562) x ABX70921	(1-4432)	
Qy	6	GlyProAlaProLeuGly-----ValArgAlaProProAlaTTPArgThrSerPro	22
Db	1135	GGTTCCCGCCCTATAGACAAGACAGTCAAGAGAGCTGGCAGGCGCTGGCGACAGGCGCG	1194
Qy	23	AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTTPValLys	42
Db	1195	CGCGCGGAGATGGGGCGCGAGCGGGCGCGCGCTGCATCCGTGCTGCTGGGTGAAG	1254
Qy	43	GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTTPArgSer	62
Db	1255	CAGCAGCGCTCGCGCGTGAAGCTGGAGCCGCGCGGCTCTGCTGCGCTGCTGGCGGAGC	1314
Qy	63	ProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIle	82
Db	1315	CCGGGGCCCGAGCGCGCGCGCGCGCGGATGCTGCTGTCCTGTATCTGAGATC	1374
Qy	83	IleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTTPGlnLys	102
Db	1375	ATCCCGCTTGAGGAAACAGACGTTCCGGGAAACATCAAGCAGTGGAAATGGCAGAA	1434
Qy	103	MetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTTP	122
Db	1435	ATGGAAAGCCCTTACGCTTTTACAGTTTCACTGTGTAAAGAGAGCAGCAGCGCACCGCTGG	1494
Qy	123	LysTTPAlaGlnValThrPheTTPCysProGluGluGlnLeuCysHisLeuTTPLeuGln	142
Db	1495	AAGTGGCGCGAGGTACTTCTGTTGTGTCAGAGAGCAGCTGTGTCACTTGTGTGCTGCAG	1554
Qy	143	ThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIle	162
Db	1555	ACCCTGGGNGATGCTGGAGAACCTGACGTCACAGCCAAAGCATTTACTTGGTATTTATC	1614
Qy	163	AsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeu	182
Db	1615	AACCCGTTTGGAGGAAAGGACAGGCAAGCGGATATATGAAAGAAAGTGGCACCACTG	1674
Qy	183	PheThrLeuAlaSerIleThrThrAspIleValThrGluHisAlaAsnGlnAlaLys	202
Db	1675	TTCACTTAGCGCTCCATCCACCTGACATCATCGTTACTGAACATGCTAATACGGCCAG	1734
Qy	203	GluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAsp	222
Db	1735	GAGACTCTGTATGAGATTACATAGACAAATACACCGGCATCGTCTGTGTGCGCGGAGAT	1794
Qy	223	GlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyVal	242
Db	1795	GGTATGTTACGCGAGGTGCTGCACCGTCTGATTGGGAGGACGACAGAGGCGCGCGGTC	1854
Qy	243	AspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIlePro	262
Db	1855	GACCAGAACCAACCCCGCGCTGCTGGTCCCCAGTAGCTCCGGATTGGAATCATTTCC	1914
Qy	263	AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer	282
Db	1915	GCAGGTCACAGACTCGGTGTGTATCTCCACCGTGGGACACGACGACGAGAAACCTCG	1974
Qy	283	AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsn	302
Db	1975	CGCGTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTGTCTCAGTCCACCACAAC	2034
Qy	303	SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIle	322
Db	2035	AGCACACTCTTCGCTACTCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATC	2094
Qy	323	LysAspSerGluLysLysArgTTPLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys	342
Db	2095	AAGACACGTGAGAAAGAACGGTGTGTGGGTCTTCCACAGATACGACTTTTTCAGGTTTAAAG	2154
Qy	343	ThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr	362

[illegible]

```
XX (CURA-) CURAGEN CORP.
PA (GETH ) GENENTECH INC.
XX
XX Rastelli L;
XX
XX WPI; 2001-514770/56.
DR P-PSDB; AAE07884.
XX
XX An isolated Sphingosine kinase polypeptide useful for treating a Sphk-
PT associated disorder especially cancer, restenosis or ischemia in a human.
PT
XX
XX Claim 8; Page 94-95; 107pp; English.
PS
XX
XX The present invention relates to sphingosine kinase (SphK) polypeptides
CC and nucleic acids encoding them. SphK is useful for treating a Sphk-
CC associated disorder especially cancers such as leukaemia, lymphoma,
CC ovarian, breast, lung, colon, testicular, stomach and skin,
CC atherosclerosis, restenosis or ischaemia and cell proliferative disease
CC or disorder associated with vascular diseases. SphK gene is used in gene
CC therapy and antisense-therapy. Sphingolipids serving as signalling
CC molecules, have recently emerged as regulators of cell growth,
CC differentiation, diverse cell phenotypes and cell death. Activation of
CC SphK by tumour necrosis factor (TNF)-alpha inhibits apoptosis in human
CC endothelial cells. The present sequence is human sphingosine kinase
CC (SphK) cDNA
XX
XX Sequence 1840 BP; 459 A; 468 C; 503 G; 410 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 9.51e-205 Length: 1840
Score: 2649.50 Matches: 495
Percent Similarity: 97.83% Conservative: 1
Best Local Similarity: 97.63% Mismatches: 0
Query Match: 87.59% Indels: 11
DB: 5 Gaps: 1

US-10-631-958-11 (1-562) x AAD14426 (1-1840)
QY 67 AlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaValGlu 86
Db 1 GCCGGCGCCCGCGCGGATGCTCTCTGTGCTGTATCTGAGATCATCGCGTTGAG 60
QY 87 GluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLysPro 106
Db 61 GAACACAGCGTTACGGGAAACATCAAGCGCAGTGGAAATGCGAGAAATGGAAGGCT 120
QY 107 TyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrpLysTrpAlaGln 126
Db 121 TACGCTTTTACAGTTTCACTGTGTAAAGAGACGACGACCGCGCTGGAAAGTGGCGCAG 180
QY 127 ValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArgGlu 146
Db 181 GTGACTTTCTGGTGTCCAGAGAGAGAGCTGTGTCTGTGTGGTGCAGACCTCGCGGAG 240
QY 147 MetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPheGly 166
Db 241 ATGCTGGAGAGCTGACGCTCCAGACCAAGCATTTACTGGTATTATCAACCCGTTTGA 300
QY 167 GlyLysGlyGlnGlyLysArgIleTrpGluArgLysValAlaProLeuPheThrIleuAla 186
Db 301 GGAAAGGACCAAGGCAAGCGGATATGAAAGAAAAGTGGCACCACCTGTTCACTTAGCC 360
QY 187 SerIleThrThrAspIleIle-----ValThr 195
Db 361 TCCATCACCTGACATCATCGTTAAACAAATTTCTATGTAACTATGTAGAGTAACTACT 420
QY 196 GluHisAlaAsnGlnAlaLysGluThrLeuTrpGluIleAsnIleAspLysTrpAspGly 215
Db 421 GAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGACGC 480
QY 216 IleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyIleGluLysArg 235
Db
```

```
Db 481 ATCTCTGTGTGCGCGAGAGATGGTATATTCTCAGCAGAGGTGCTGCACGGTCTGATTGGGAGG 540
QY 236 ThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSer 255
Db 541 ACCGAGAGGAGCGCGGGGTCGACACAGAACACACCCCGGGCTGTGCTGGTCCCCAGTAGC 600
QY 256 LeuArgIleGlyIleIleProIleProIleGlySerThrAspCysValCysTrpSerThrValGly 275
Db 601 CTCGGATTGGAATCATTTCCCGCAGGCTCAACGACCTGCGTGTGTATTCTCCACCGTGGC 660
QY 276 ThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAsp 295
Db 661 ACCAGACGACGAGAAACCTCGCGCTGCATATCTTGTGGGGACTCGCTGGCCATGGAT 720
QY 296 ValSerSerValHisHisAsnSerThrLeuLeuArgTrpSerValSerLeuLeuGlyTrp 315
Db 721 GTGTCTCAGTCCACCAACACAGCACACTCTTTCGCTACTCCGTGTCCCTGCTGGGTAC 780
QY 316 GlyPheTrpGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArg 335
Db 781 GGCTTCTACGGGACATCATCAGGACAGTGAAGAAACGCTGTTGGGTCTTGGCCAGA 840
QY 336 TyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTrpGluGlyThrValSer 355
Db 841 TACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAGAGGACAGTGTCC 900
QY 356 PheLeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGly 375
Db 901 TTCTCTCTGCACACACACGCTGGGATCTCCAAGGATAGGAAGCCCTGCCGGCAGGA 960
QY 376 CysPheValCysArgGlnSerLysGlnGlnLeuGluGlnLysLysAlaLeuTrp 395
Db 961 TGCTTTTTCGAGGCAAAAGCAAGCAGCAGCTGGAGGAGGAGCAGAGAAAGCACTGTAT 1020
QY 396 GlyLeuGluAlaAlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAla 415
Db 1021 GGTTTGGAAGCTCGGAGGAGCTGGAGAGTGGCAAGTGTCTGTGGAAAGTTTCTGGCC 1080
QY 416 IleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAla 435
Db 1081 ATCAATGCCACAAACATGTCTGTGCTTGTTCGCCGAGCCCCAGGGGCTCTCCCCGCT 1140
QY 436 AlaHisLeuGlyAspGlySerSerAspLeuIleLeuLysCysSerArgPheAsn 455
Db 1141 GCCCAGCTGGGAGACGGGTCTTCTGACCTCATCTCATCCGAAAATGTCAGGTTCAAT 1200
QY 456 PheLeuArgPheLeuIleArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheVal 475
Db 1201 TTTCTGAGATTCTCATCAGGCACACCAACAGCAGGACCAAGTTTGACTTTCACTTTGT 1260
QY 476 GluValTrpArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSer 495
Db 1261 GAAGTTTATCGCTCAAGAAATTCAGTTTACGTTCGAAGCACATGGAGGATGAGGACAGC 1320
QY 496 AspLeuLysGluGlyLysLysLysArgPheGlyHisIleCysSerSerHisProSerCys 515
Db 1321 GACCTCAAGAGGGGGGAAAGCGCTTTTGGGCACATTTTGCAGCAGCAGCCCTCTCTGC 1380
QY 516 CysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAla 535
Db 1381 TGCTGCACCGTCTCCACACAGCTCTTGGAACTGCCACCGGGAGGTCCTGCAGAGCCCTGCC 1440
QY 536 IleGluValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsn 555
Db 1441 ATCAGGTCAGGTCCTGCTGCCAGCTGTTGCACTCTTTGGCAGGAAATTTGAAGAGAT 1500
QY 556 ProLysProAspSerHisSer 562
Db 1501 CCGAAGCAGACTCACACAGC 1521

RESULT 9
AAAS0510
ID AAAS0510 standard; cDNA; 4231 BP.
```

XX AAA50510;  
 AC  
 XX  
 DT 05-DEC-2000 (first entry)  
 XX  
 DE Human sphingosine kinase C cDNA.  
 XX  
 KW Sphingosine kinase C; SKC; human; drug screening; infection;  
 KW antiinflammatory; antiallergic; anticancer; inflammation; allergy;  
 KW cancer; therapy; diagnosis; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 71..1453  
 FT /\*tag= a  
 XX  
 XX WO200052173-A2.  
 PN  
 XX  
 PD 08-SEP-2000.  
 XX  
 XX 02-MAR-2000; 2000WO-CA000223.  
 PF  
 XX 02-MAR-1999; 99US-0122516P.  
 PR  
 XX (ALLX ) NPS ALLELIX CORP.  
 PA  
 XX Munroe D, Gupta A, Falzone GR;  
 FI  
 XX WPI; 2000-572185/53.  
 DR  
 DR P-PSDB; AAY96059.  
 XX  
 PT New human sphingosine kinase A, B and C polynucleotides and polypeptides  
 PT useful in e.g. chromosome and gene mapping, and detecting inflammation or  
 PT disease associated with abnormal levels of sphingosine kinase expression.  
 XX  
 XX Disclosure; Fig 7; 81pp; English.  
 PS  
 XX The present sequence is that of an isolated polynucleotide encoding human  
 CC sphingosine kinase C (SKC, see AAY96059), an enzyme that phosphorylates  
 CC sphingosine to form sphingosine 1-phosphate. The polynucleotide was  
 CC isolated from an HeLa cDNA library by PCR amplification. The invention  
 CC provides polynucleotides (see AAA50508-10) and polypeptides (see AAY96057  
 CC -59) for the human sphingosine kinase (SK) homologues SKA, SKB and SKC.  
 CC The polynucleotides may be used as hybridization probes, in the  
 CC construction of PCR primers for chromosome and gene mapping, in the  
 CC recombinant production of SKA, SKB and SKC, and in the generation of  
 CC antisense DNA or RNA. They can be used to detect inflammation or disease  
 CC associated with abnormal levels of SK expression, or to detect  
 CC differences in gene sequence between normal and carrier or affected  
 CC individuals. Host cells expressing SK can be used in drug screening.  
 CC Human SK specific antibodies, inhibitors, ligands or their analogues are  
 CC useful as bioactive agents to treat inflammation or disease including  
 CC viral, bacterial or fungal infections, allergic responses, mechanical  
 CC injury associated with trauma, hereditary diseases, lymphoma or  
 CC carcinoma, and other conditions with activate the genes of kidney, lung,  
 CC heart, lymphoid or tissues of the nervous system  
 XX  
 SQ Sequence 4231 BP; 1022 A; 1021 C; 1086 G; 1067 T; 0 U; 35 Other;  
 Alignment Scores:  
 Pred. No.: 3,32e-199 Length: 4231  
 Score: 2587.00 Matches: 481  
 Percent Similarity: 99.79% Conservative: 1  
 Best Local Similarity: 99.59% Mismatches: 1  
 Query Match: 85.52% Indels: 0  
 DB: 3 Gaps: 0  
 US-10-631-958-11 (1-562) x AAA50510 (1-4231)  
 Qy 80 SerGluIleAlaValGluThrAspValHisGlyLysHisGlnGlySerGlyLys 99  
 Db 2 TCTGAGATCATCGCCGTTGAGGAACAGACGTTTACGGGAACATCAAGGCAGTGGAAAA 61

100 TrpGlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgArg 119  
 62 TGGCAGAAAAATGGAAAAAGCCTTACGCTTTTACAGTTTCACTGTGTAAAGAGACGACGG 121  
 120 HisArgTyrLysTrpAlaGlnValThrPheTyrCysProGluGluGlnLeuLysHisLeu 139  
 122 CACCGCTGGAAAGTGGGGCGCAGGTGACTTTCTGGTGTCCAGAGGAGCAGCTGTGTCTACTG 181  
 140 TrpLeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeu 159  
 182 TGGCTGCAGACCCCTGCGGAGATGCTCGAGAAGCTGACGTCAGACCAAAAGCATTTACTG 241  
 160 ValPheIleAsnProPheGlyGlyLysGlnClyLysArgIleTyrGluArgLysVal 179  
 242 GTATTATTATCAACCCGTTTGGAGGAAAAAGGCAAGCGAAGCGGATATATGAAGAAAAAGTG 301  
 180 AlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsn 199  
 302 GCACCACTGTTACCTTAGCTCCATCCACTGCATCATCGTTACTGGAACATGCTAAT 361  
 200 GlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysVal 219  
 362 CAGGCCAAGGAGACTCTGTATGAGATTAAACATACAAATACGACGCGCATCGTCTGTGTC 421  
 220 GlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSer 239  
 422 GCGCGAGATGGTATGTTTCAGCGAGGTGCTGCACCGTCTGATTGGGAGGACGACAGGAGC 481  
 240 AlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGly 259  
 482 GCCGGGGTGCACCAAGAACCCACCCCGGGCTGTGTGTCCTCCAGTAGCTCCGGATTGGA 541  
 260 IleIleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAla 279  
 542 ATCATTCGCGCAGGGTCAACGAGCTGCGTGTGTTACTCCACCGTGGGCACGACGACGCA 601  
 280 GluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerVal 299  
 602 GAAACCTCGCGCGTGCATATCGTTGTTGGGACCTCGCTGGCCATGGATGTCTCTCAGTC 661  
 300 HisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGly 319  
 662 CACCACAAACAGCACACTCTTCGCTACTTCGCTGCTCCCTGCTGGGCTACGGCTTCTACGGG 721  
 320 AspIleIleLysAspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSer 339  
 722 GACATCATCAAGGACAGTGAAGAAACGGTGTGGTCTTGGCCAGATACGACTTTTCA 781  
 340 GlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAla 359  
 782 GGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTCTCTCCCTCCCTGCA 841  
 360 GlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCys 379  
 842 CAACACACGGTGGGATCTCCAAGGGATAGGAAGCCCTGCCGGGAGGAGTCTTGTGTTGC 901  
 380 ArgGlnSerLysGlnGlnLeuGluGlnLysAlaLeuLysValLeuTyrGlyLeuGluAla 399  
 902 AGGCAAGCAAGCAGCAGCTGGAGGAGGAGCAGAAAGAACGACCTGTATGGTTTGGAGCT 961  
 400 AlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThr 419  
 962 GCGGAGGACGCTGGAGGAGTGGCAAGTGTCTGTGGGAAGTTTCTGGGCAATCAATGCCACA 1021  
 420 AsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGly 439  
 1022 AACATGTCCTGTGCTTGTGCGCGGAGCCCGCAGGGGCTCTCCCGGCTGCCCACTTGGGA 1081  
 440 AspGlySerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPhe 459  
 1082 GACGGGCTTCTTGACCTCATCTCTCCGNNATGCTCCAGGTTCAATTTTCTGAGATT 1141

Qy 460 LeuLeuArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArg 479  
 Db 1142 CTCATCAGGCACACCAACAGCAGCAGCAGTTTGACTTCACTTTTGTGAAGTTTATCGC 1201  
 Qy 480 VallysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGlu 499  
 Db 1202 GTCAAGAAATTCAGTTTACGTCTGAAGCACATGAGGATGAGGACAGCAGCTCAAGGAG 1261  
 Qy 500 GlyGlyLysLysArgPheGlyHisLeuLysSerHisProSerCysCysCysThrVal 519  
 Db 1262 GGGGGGAAGAGCCCTTTGGCARTTTGACGACGACCCCTCTCTGCTGCACCGTC 1321  
 Qy 520 SerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaLeuValArg 539  
 Db 1322 TCCAACAGCTCTGGAACCTGACGCGGAGGCTCTGCACAGCCCTGCCATCGAGTCA 1381  
 Qy 540 ValHisCysGlnLeuValArgLeuPheAlaArgGlyLeuGluGluAsnProLysProAsp 559  
 Db 1382 GTCCACTGCCAGCTGGTTGCAGCTCTTTGCACGAGAAATTGGAAGAGAATCCGAAGCCAGAC 1441  
 Qy 560 SerHisSer 562  
 Db 1442 TCACACAGC 1450  
 RESULT 10  
 ID ADS10370 standard; DNA; 4702 BP.  
 AC ADS10370;  
 XX  
 XX 16-DEC-2004 (first entry)  
 XX Human therapeutic DNA - SEQ ID 607.  
 XX  
 XX antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnerary;  
 KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;  
 KW aplastic anaemia; cancer; wound healing; gene therapy; ds; gene.  
 XX  
 XX Homo sapiens.  
 XX  
 XX WO2004080148-A2.  
 XX  
 XX 23-SEP-2004.  
 XX  
 XX 30-SEP-2003; 2003WO-US030720.  
 XX  
 XX 02-OCT-2002; 2002US-0416186P.  
 PR  
 XX (NUVE-) NUVELO INC.  
 XX  
 XX Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y;  
 PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;  
 DR WPI; 2004-668857/65.  
 DR P-PSDB; ADS11054.  
 XX  
 XX New polynucleotide, useful in preparing a composition for diagnosing or  
 PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,  
 PT aplastic anemia or cancer for promoting wound healing.  
 XX  
 XX Claim 1; SEQ ID NO 607; 718pp; English.  
 XX  
 XX The invention relates to a novel isolated polynucleotide and the encoded  
 CC polypeptide. The molecules of the invention demonstrate antiinflammatory,  
 CC neuroprotective, antianaemic, cytostatic and vulnerary activities and may  
 CC be useful in preparing a composition for diagnosing or treating  
 CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell  
 CC disorders, such as aplastic anaemia or cancer, as well as for promoting  
 CC wound healing. The molecules may also be utilised during gene therapy  
 CC procedures. The current sequence is that of a human therapeutic DNA of  
 CC the invention. The current sequence is not shown explicitly within the  
 CC specification but can be accessed from the WIPO web-site.

SQ Sequence 4702 BP; 970 A; 1226 C; 1364 G; 1118 T; 0 U; 24 Other;  
 Alignment Scores:  
 Pred. No.: 1,86e-195 Length: 4702  
 Score: 2541.50 Matches: 505  
 Percent Similarity: 65.72% Conservative: 3  
 Best Local Similarity: 65.33% Mismatches: 7  
 Query Match: 84.02% Indels: 259  
 DB: 13 Gaps: 4  
 US-10-631-958-11 (1-562) x ADS10370 (1-4702)  
 Qy 6 GlyProAlaProLeuGly-----ValArgAlaProProAlaTrpArgThrSerPro 22  
 Db 749 GGTTCCCGCCCTATAGACAAGACAGTCACAAGGAGCTGGCAGGCTGGCGGAGCAGCCG 808  
 Qy 23 AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLys 42  
 Db 809 GCGGCGAGATGGGGCGACGCGGGGCGGCGGCGCTGCCAATCCGTCGTGTGGGTGAAG 868  
 Qy 43 GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSer 62  
 Db 869 CAGCAGCGCTGCGCGCTGAGCCTGGAGCCCGCGGGGCTCTGTGCGCTGGTGGCGGAGC 928  
 Qy 63 ProGlyProGlyAlaGlyAlaProGlyAla----- 72  
 Db 929 CCGGGCCCGGAGCGCGCGCCCGCGCGC-GGCTCACCAAGGCCAGCCCTCCTTCCCA 987  
 Qy 72 ----- 72  
 Db 988 GACTCACCTAGCACTGGTTATTTTGCAGTATGTTCCAAAGACTGGGCAGCGCTTACACCTT 1047  
 Qy 72 ----- 72  
 Db 1048 CTGTTTAGCTGGTTTCTTGCGCACTGCTCCAGCAGGAAACGGCGCTGGTACTGCCAGGG 1107  
 Qy 72 ----- 72  
 Db 1108 CCTGCACCCACCCCTGCTGCCCATAGCTCAGACGTCCTCCGAGGCCAGAGCTGGAGATAA 1167  
 Qy 72 ----- 72  
 Db 1168 CCCAGCTCCAGCTGCTGCTGCTGCTGCGCTCTCCAGCCAGCGGCGCTGCTGCTGT 1227  
 Qy 72 ----- 72  
 Db 1228 GGATCTCTGTGTGTCGCCGGTTCACTGCTGCTGTTCTGTATCTTGTCTCCAGAGCTATCAGTC 1287  
 Qy 72 ----- 72  
 Db 1288 ATAGCGGACGGGAGGCTGATGCAGGTTTCCCGCGACCTGCACAGCATGGAGTGGGACGG 1347  
 Qy 73 -----AspAlaCysSer 76  
 Db 1348 CCGGTGTTCTGTGTGATGATGCCAGCTGAAGAGCGTGGTGTGTTTACAGATGCTGCTCT 1407  
 Qy 77 ValProValSerGluLeuLeuAlaValGluGluThrAspValHisGlyLysHisGlnGly 96  
 Db 1408 GTGCTGTATCTAGATCATCGCGTTGAGAAACACAGCTTCAACGGGAAACATCAAGGC 1467  
 Qy 97 SerGlyLysTrpGlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArg 116  
 Db 1468 AGTGGAAATGGCAGAAAATGGAAAAGCCCTTACGTTTTTACAGTCTCACTGTGTAAGAGA 1527  
 Qy 117 AlaArgHisArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeu 136  
 Db 1528 GCACGACGGCAGCGCTGGAAGTGGCGCAGGTCTTCTGGTGTCCAGAGGACGAGCTG 1587  
 Qy 137 CysHisLeuTrpLeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLys 156  
 Db 1588 TGTCACTGTGGCTGCAGACCCCTGCGGAGATGCTGGAGAGCTG----- 1632  
 Qy 157 HisLeuLeuValPheIleAsnProPheGlyLysGlyGlnGlyLysArgIleTyrGlu 176



Db 1632 ----- 1632  
Qy 177 ArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGlu 196  
Db 1633 -----ATTACTGAA 1641  
Qy 197 HisAlaAsnGlnAlaIysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIle 216  
Db 1642 CATGCTAATCAGGCCAAGAGACTCTGTATGAGATTAAACATAGACAAATACGACGGCATC 1701  
Qy 217 ValCysValGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyValArgThr 236  
Db 1702 GTCTGTGTGGCGGAGATGGTATGTTTCAGCGAGGTGTGCACGGTCTGATTGGGAGGAGC 1761  
Qy 237 GlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeu 256  
Db 1762 CAGAGGAGCGCGGGGTGCACAGAAACACCCCGGGGTGTGTGTCCTCCAGTAGCCTC 1821  
Qy 257 ArgIleGlyIleIleProAla----- 263  
Db 1822 CGGATTGGAATCATTTCCCGCAGGCTTTCGAAATGACCTGGCAGGGGACGAGGTGTCTGTC 1881  
Qy 263 ----- 263  
Db 1882 CTCTCTGGCCCTGTGTCTGGCCCCGAGGGTGGCGGCATGTGTGCACACTTTCACCTCTGTC 1941  
Qy 263 ----- 263  
Db 1942 ACAGCTCTGGATGTGAGCACCGCAGTCATCCCATTTTATGGATGAAGACAGGAGGACT 2001  
Qy 264 -----GlySerThrAspCysVal 269  
Db 2002 GGGGAGCATGTGGGCCCCGGTGAGAACCGTGTGTGGCTTGGACGGGTCAACGGACTCGCTG 2061  
Qy 270 CysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGly 289  
Db 2062 TGTTACTCACCGTGGGACACGACGACGAGAAACCTCGCGCTGCATGTGTGTGG 2121  
Qy 290 AspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeuLeuArgTyrSer 309  
Db 2122 GACTCGCTGGCCATGGATGTGTCTCAGTCAGTCACACACACAGACACTCTTCGCTACTCC 2181  
Qy 310 ValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGluLysIysArg 329  
Db 2182 GTGTCCCTCTGGGCTACGGCTTCTACGGGACATCATCAGGACAGTCAGAAAGACGG 2241  
Qy 330 TrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisCys 349  
Db 2242 TGGTTGGGCTTGGCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGC 2301  
Qy 350 TyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArg 369  
Db 2302 TATGAAGGACAGTGTCTCTCTCCCTGCACACACAGGTGGGATCTCCAAGGATAGG 2361  
Qy 370 LysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLeuGluGlu 389  
Db 2362 AAGCCCTGCGGGCAGGATGCTTTGTTTTCAGGCAAGCAAGCAGCAGCTGGAGGAGGAG 2421  
Qy 390 GlnLysLysAlaLeuTyrGlyLeuGluAlaGluAspValGluGluTrpGlnValVal 409  
Db 2422 CAGAAGAAGCACTGTATGGTTTGAAGCTGCGGAGGACGTGGAGGATGGCAAGTCTGTC 2481  
Qy 410 CysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerPro 429  
Db 2482 TGTGGAGTTTCTGGCATCATGCCACAAACATGTCTGTGTGTGTGGCGGAGCCCC 2541  
Qy 430 ArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuLeuLeuArg 449  
Db 2542 AGGGGCTCTCCCGGCTGCCACTTGGAGACGGGTCTTCTGACCTCATCTCATCCGG 2601  
Qy 450 LysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsnGlnGlnAspGln 469

Db 2602 AAATGCTCCAGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAACGACGAGACCAG 2661  
Qy 470 PheAspPheThrPheValGluValTyrArgValLysLysPheGlnPheThrSerLysHis 489  
Db 2662 TTTGACTTCACTTTGTTGAAGTTTATCGCTCAAGAAATTTCCAGTTTACGTCGAAGCAC 2721  
Qy 490 MetGluAspGluAspSerAspLeuLysGluGlyGlyLysLysArgPheGlyHisIleCys 509  
Db 2722 ATGGAGATGAGACGACCTCAAGGAGGGGGGAGAAAGCGCTTTTGGGCACATTTGC 2781  
Qy 510 SerSerHisProSerCysCysCysThrValSerAsnSerSerTrpAsnCysAspGlyGlu 529  
Db 2782 AGCAGCACCCCTCTCTGCTGTGCACCGCTCTCCACAGCTCTGGAACCTGCACGGGAG 2841  
Qy 530 ValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuValArgLeuPheAla 549  
Db 2842 GTCTGCACAGCCCTGCATCGAGTCAAGTCCAGTCCAGCTGGTGTTCGACTCTTTGCA 2901  
Qy 550 ArgGlyIleGluGluAsnProLysProAspSerHisSer 562  
Db 2902 CGAGGAATTGAAGAGAAATCCGAGGCCAGACTCACACAGC 2940  
RESULT 11  
AAS77728  
ID AAS77728 standard; cDNA; 2241 BP.  
XX  
AC AAS77728;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #13532.  
XX  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200175067-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US008631.  
XX  
PR 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
XX  
XX P-PSDB; ABG13541.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
XX biodiversity.  
XX  
XX Claim 1; SEQ ID NO 13532; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in



CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 2241 BP; 453 A; 627 C; 700 G; 461 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 2,03e-184 Length: 2241  
 Score: 2400.00 Matches: 483  
 Percent Similarity: 64.39% Conservative: 7  
 Best Local Similarity: 63.47% Mismatches: 14  
 Query Match: 79.34% Indels: 258  
 DB: 5 Gaps: 6

US-10-631-958-11 (1-562) x AAS77728 (1-2241)

Qy 6 GlyProAlaProLeuGly-----ValArgAlaProProAlaTTPArgThrSerPro 22  
 Db 14 GGTTCGCGCCCTATAGACAGACAGTCAACAGGAGCTGGCAGCGCTGGCGACAGGCCG 73  
 Qy 23 AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTTPVallys 42  
 Db 74 GCGCGGAGATGGGGCGACGCGGGCGCGAGCGCTGCANTCCGTGCTGGGTGAAG 133  
 Qy 43 GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTTPArgSer 62  
 Db 134 CAGCAGCGCTGCGCGCTGAGCGTGGAGCCCGCGCGGCTCTGCTGCGTGGTGGCGAGC 193  
 Qy 63 ProGlyProGlyAlaGlyAlaProGlyAla----- 72  
 Db 194 CCGGGCGCGGAGCG 252  
 Qy 72 ----- 72  
 Db 253 GACTCAGTACCTGGTTATTTTGCAGTATGTTCCAGACTGGGCGAGCGTTTACACCTT 312  
 Qy 72 ----- 72  
 Db 313 CTGTTAGCTGTTCTCTGCACTGTCTCCAGCAGGGAACGGCGCTGGTACTGCCAGGG 372  
 Qy 72 ----- 72  
 Db 373 CCTGCAACCCCTGTGTCCTCCATAGCTCAGACGTCCCGAGGCCAGAGCTGGAGATAA 432  
 Qy 72 ----- 72  
 Db 433 CCCAGCTCCCAGCTGCTGACTCTGCACGCTGGCCCTCTCCAGCCAGCGGGCTGTGCTGT 492  
 Qy 72 ----- 72  
 Db 493 GGATCTCTGTGTCCCGGTTCACTGCGCTGTTCTTGTTATCTTGTCCAGAGCTATCAGTC 552  
 Qy 72 ----- 72  
 Db 553 ATAGCGGAGCGGAGGCTGATGCAGGTTTCCGCGACCTGACAGCATGGAGTGGGACGG 612  
 Qy 73 ----- AspAlaCysSer 76  
 Db 613 CCGGTGTTCTGTGTGATGCCAGCTGAAGGACGTGGTGTGTTTACAGATGCTGCTCT 672  
 Qy 77 ValProValSerGluIleAlaValGluThrAspValHisGlyLysHisGlnGly 96  
 Db 673 GTGCTGTATCTGAGATCATCGCGTGTGAGGAAACAGACGTTTCCCGGAAACATCAAGGC 732  
 Qy 97 SerGlyLysTTPGlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArg 116  
 Db 733 AGTGGAAATGGCAGAAATGGAAAGCCCTTACGCTTTTACAGTTCACTGTGTAAAGAGA 792

Qy 117 AlaArgArgHisArgTTPLysTTPAlaGlnValThrPheTTPCysProGluGluGlnLeu 136  
 Db 793 GCACGACGGCACCGCTGGAGTGGCGCAGGTGACTTCTTGGTGTCCAGAGGACGACGTG 852  
 Qy 137 CysHisLeuTTPLeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLys 156  
 Db 853 TGTCACTTGTGGTGCAGACCCCTGCGGAGATGCTGGAGAAGCTG----- 897  
 Qy 157 HisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGlu 176  
 Db 897 ----- 897  
 Qy 177 ArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGlu 196  
 Db 898 -----ATTACTGAA 906  
 Qy 197 HisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIle 216  
 Db 907 CATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGACGGCATC 966  
 Qy 217 ValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyValThr 236  
 Db 967 GTCTGTGTCGCGGAGATGTTATGTTCCAGCAGGTGCTGCACGGTCTGATTTGGAGAGCG 1026  
 Qy 237 GlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeu 256  
 Db 1027 CAGAGGAGCGCGGGTCCGACAGAACCCCGCGGCTGTGCTGGTCCCCCAGTAGCCTC 1086  
 Qy 257 ArgIleGlyIleIleProAla----- 263  
 Db 1087 CGGATTGGAATCATTCCTCCGACAGACATGTGGGCCCCCGTGAGAACGCTGGTGGCTTGGAC 1146  
 Qy 264 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 283  
 Db 1147 GGGTCAACGAGACTGCGTGTGTTACTCCACCCTGGGACACGACGACGAGAAACCTCGCG 1206  
 Qy 284 LeuHisIleValVal----- 288  
 Db 1207 CTGCATATCGTTGTGGCTGCTGCTCCGCGAGGCTCGGAACACCACCGCATCCCGCCATACT 1266  
 Qy 288 ----- 288  
 Db 1267 GCCTGTGGCAGTGGGCACGGGACAGCTGTGCTGGGCTGCCAGCTGTGGAACGCATGCCTC 1326  
 Qy 289 -----GlyAspSerLeuAlaMetAsp 295  
 Db 1327 TGTGAGCCCTCGAGGCTTCAGTCCAGATGTCAGAGCCCCCGGGAGCTCGCTGGCCATGGAT 1386  
 Qy 296 ValSerSerValHisAsnSerThrLeuLeuArgTyrSerValSerLeuGlyTyr 315  
 Db 1387 GTGTCTCTCAGTCCACCAACACAGCACACTCTTCGCTACTCCGTGTCCTGCTGGGCTAC 1446  
 Qy 316 GlyPheTyrGlyAspIleIleLysAspSerGluLysLysArgTTPLeuGlyLeuAlaArg 335  
 Db 1447 GGCTTCTACCGGGACATCATCAAGGACAGTCAGAGAAGAACGGTGGTGGGTCTTGGCAGA 1506  
 Qy 336 TyrAspPheSerGlyLysThrPheLeuSerHisCysTyrGluGlyThrValSer 355  
 Db 1507 TACGACITTTTCAGTATAAAGACCTTCTCTCCACCACCTGCTATAGAAGGACAGTGTCC 1566  
 Qy 356 PheLeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGly 375  
 Db 1567 TTCTCTCTCCACAAACACCGTGGGATCTCCAAGGAATAGGAAGCCCTGCGGGCAGGA 1626  
 Qy 376 CysPheValCysArgGlnSerLysGlnLeuGluGlnLysLysAlaLeuTyr 395  
 Db 1627 TGCTTTGTCAGGCAAGCAAGCAGCAGCTGGAGGAGGAGGAGAAAGAAAGCACTGTAT 1686  
 Qy 396 GlyLeuGluAlaAlaGluAspValGluGluTTPGlnValValCysGlyLysPheLeuAla 415  
 Db 1687 GGTTTGGAAGCTCGGAGGACGTGGAGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGGCC 1746  
 Qy 416 IleAsnAlaThrAsnMetSerCysAlaCysArgSerProArgGlyLeuSerProAla 435

Db	1747		ATCAATGCCAACAATGCTCTGTGCTGTGTCGCCGAGCCCAAGGGCCCTCTCCCGGCT	1806
QY	436		AlaHisLeuGlyAspGlySerSerAspLeuileLeuileArgLysCysSerArgPheAsn	455
Db	1807		GCCCACTGGGAGACGGGTCTTCGACTCATCTCATCCGGAATGCTCAGGTTCAAT	1866
QY	456		PheLeuArgPheLeuileArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheVal	475
Db	1867		TTTCTGAGATTCTCATCAGGCACACCAACCAGCAGCACAGTTTGACTTCACTTTGTT	1926
QY	476		GluValTyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSer	495
Db	1927		GAAATTATCGCGTCAAGAATTCTCTGTTTACGTGGAAGCACATGGAAGATGAGGACAGC	1986
QY	496		AspLeuLysGluGlyGlyLysLysArgPheGlyHisIleCysSerSerHisProSerCys	515
Db	1987		GACCTCAAGAGAGGGGGGAGAGCGCTTTGGGCACATTTGCAGCAGCCACCCCTCTCTGC	2046
QY	516		CysCysThrValSerAsnSerSerTyrAsnCysAspGlyGluValLeuHisSerProAla	535
Db	2047		TGCTGCACCGTCTCCACAGCTCTGGAACTGCGACGGGAGGTCTCTGCACAGCCCTGCC	2106
QY	536		IleGluValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsn	555
Db	2107		ATCAGGTCAGG-----GCGGTGGTGATGGCAAGAGGCTGTGCAAAAGGTC	2151
QY	556	Pro	556	
Db	2152	CCA	2154	
RESULT 12				
ACC70838				
ID	ACC70838	standard; DNA; 3975 BP.		
XX	XX			
AC	ACC70838;			
XX	XX			
XX	20-NOV-2003	(first entry)		
XX	Human Sphingosine kinase 4 coding sequence.			
DE	DE			
XX	XX			
KW	KW	Human; enzyme; haemostatic; sphingosine kinase 4; SPHK4;		
KW	KW	platelet transfusion; platelet stabiliser; gene; ds.		
XX	XX			
OS	OS	Homo sapiens.		
XX	XX			
FH	Key	Location/Qualifiers		
FT	CDS	1..1251		
FT	FT	/*tag= a		
FT	FT	/product= "Sphingosine kinase 4"		
XX	XX			
PN	WO2003031627-A1.			
XX	XX			
PD	17-APR-2003.			
XX	XX			
PF	28-SEP-2001; 2001WO-JP008537.			
XX	XX			
PR	28-SEP-2001; 2001WO-JP008537.			
XX	XX			
PA	(HOKK-) HOKKAIDO TECHNOLOGY LICENSING OFFICE CO.			
PA	(CHBI-) CHEM BIOLOGY INST.			
XX	XX			
PI	Igarashi Y, Kihara A;			
XX	XX			
DR	WPI; 2003-354917/33.			
DR	P-PSDB; ABR56301.			
XX	XX			
PT	Platelet derived polypeptides with sphingosine kinase activity for			
PT	treatment of sphingosine related disorders.			
XX	XX			
PS	Claim 4; Page 30-31; 39pp; Japanese.			
XX	XX			
CC	The present sequence is the coding sequence for human sphingosine kinase			

CC	4	(SPHK4). The kinase can be used for the diagnosis and treatment of sphingosine related disorders. The kinase can also be potentially used for controlling toxicity of platelet transfusion and as a platelet stabiliser	XX	
CC	50	Sequence 3975 BP; 943 A; 978 C; 1042 G; 1012 T; 0 U; 0 Other;	XX	
CC		Alignment Scores:	XX	
CC		Pred. No.:	XX	
CC		Score:	XX	
CC		Percent Similarity:	XX	
CC		Best Local Similarity:	XX	
CC		Query Match:	XX	
CC		DB:	XX	
CC		US-10-631-958-11 (1-562) x ACC70838 (1-3975)	XX	
CC	Qy	147 MetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPheGly	3975	166
CC	Db	1 ATGCTGGAGAAGCTGAGCTGCAGACCACCAAGCATTTACTTGGTATTTATCAACCGTTTGA	416	60
CC	Qy	167 GlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeuAla	186	186
CC	Db	61 GGAAGAGGACAAAGCAAGCGGATATATGAAGAAAGTGGCACCACTGTTCACTTAGCC	120	120
CC	Qy	187 SerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeuTyr	206	206
CC	Db	121 TCCATCACCTGACATCATCTGTTACTGAACATGCTAATCAGGCCAAGGAGACTCTGTAT	180	180
CC	Qy	207 GluIleAsnIleAspLysThrAspGlyIleValCysValGlyGlyAspGlyMetPheSer	226	226
CC	Db	181 GAGATTAACTAGACAAATACAGCGGCATCTGTGTGCGCGAGATGTTATGTTACG	240	240
CC	Qy	227 GluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHis	246	246
CC	Db	241 GAGGTGTCGACGGTCTGATTGGGAGGACGACAGAGCGCGGGGTTCACCAACACAC	300	300
CC	Qy	247 ProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySerThr	266	266
CC	Db	301 CCCCGGGCTGTGCTGGTCCCAGTAGCCTCCGGATTTGGAATCATTTCCGCGAGGGTCAACG	360	360
CC	Qy	267 AspCysValCysThrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIle	286	286
CC	Db	361 GACTCGGTGTGTTACTCCACCGTGGGCACACGACGACGAGAAACCTCGCGCTGATATC	420	420
CC	Qy	287 ValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeuLeu	306	306
CC	Db	421 GTTGTGGGGACTCGCTGGCCATGGATGTGTCTCAGTCCAGTCCACCACACAGCACACTCTCT	480	480
CC	Qy	307 ArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleAspSerGlu	326	326
CC	Db	481 CGCTACTCCGGTGTCCCTGTCTGGGCTACGGCTTCTACGGGGACATCATCAAGGACAGTGA	540	540
CC	Qy	327 LysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSer	346	346
CC	Db	541 AAGAAACGGTGTGGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGACCTTCCTCTCC	600	600
CC	Qy	347 HisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerPro	366	366
CC	Db	601 CACCACCTGCTATGAAGGGACAGTGTCTTCTCTCCCTGCACAAACACACGCTGGGATCTCCA	660	660
CC	Qy	367 ArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnLeu	386	386
CC	Db	661 AGGGATAGGAAGCCCTCCGGGGCAGGATGCTTTGTTTGCAGGCAAGACGAGAGCTG	720	720
CC	Qy	387 GluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGluTrp	406	406
CC	Db	721 GAGGAGGAGCAGAAAGACACTGTATGTTTGGAAAGCTCGGAGGACGTGGAGAGTGG	780	780
CC	Qy	407 GlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArg	426	426
CC	Db	781 CAAGTCGTCTGTGGGAAGTTCTTGCCATCAATGCACAAACATCTCTGTTGTTGTGCG	840	840

The present sequence is the coding sequence for human sphingosine kinase

Qy 427 ArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuLeu 446  
 Db 841 CGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCGACCTCATC 900  
 Qy 447 LeuLeuArgLysCysSerArgPheAsnPheLeuArgPheLeuLeuArgHisThrAsnGln 466  
 Db 901 CTCTATCCGGAATGCTCCAGGTTCATTTCTGAGATTTCTCATCAGGCACACCAACCAG 960  
 Qy 467 GlnAspGlnPheAspPheThrPheValGluValTrpArgValLysLysPheGlnPheThr 486  
 Db 961 CAGGACCATTTGACTTCACTTTTGTGAAGTTTATCGCGTCAAGAATTCAGATTACG 1020  
 Qy 487 SerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysLysArgPheGly 506  
 Db 1021 TCGAAGCACATGGAGGATGAGGACGACCTCAAGGAGGGGGGGAAGCGCTTTGGG 1080  
 Qy 507 HisLeuCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsnCys 526  
 Db 1081 CACATTTGCAGCAGCCACCCCTCGCTGTGTCACCGTCTCCAAACAGCTCTGGAACATGC 1140  
 Qy 527 AspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuValArg 546  
 Db 1141 GACGGGAGGTCCTGCACAGCCCTGCCATCGAGGTCAGAGTCCACTGCCAGCTGGTTGGA 1200  
 Qy 547 LeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 562  
 Db 1201 CTCTTTGCAGGAGAAATTGAAGAGATCCGAAAGCCAGACTCACACAGC 1248

## RESULT 13

AAS77730

ID AAS77730 standard; cDNA; 2186 BP.

XX AAS77730;  
 AC  
 XX

DT 13-FEB-2002 (first entry)  
 XX

DE DNA encoding novel human diagnostic protein #13534.  
 XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 XX food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 OS Homo sapiens.

XX WO200175067-A2.  
 XX

PN 11-OCT-2001.  
 XX

PD 30-MAR-2001; 2001WO-US008631.  
 XX

PF 31-MAR-2000; 2000US-00540217.  
 XX

PR 23-AUG-2000; 2000US-00649167.  
 XX

XX (HYSE-) HYSEQ INC.  
 PA

XX Drmanac RT, Liu C, Tang YT;  
 PI

XX WPI; 2001-639362/73.  
 DR

XX P-PSDB; ABG13543.  
 DR

XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensic, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.

XX Claim 1; SEQ ID NO 13534; 103pp; English.  
 PS

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (II) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is

CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensic, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 2186 BP; 455 A; 637 C; 669 G; 425 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 6.81e-129 Length: 2186  
 Score: 1714.50 Matches: 345  
 Percent Similarity: 63.65% Conservative: 0  
 Best Local Similarity: 63.65% Mismatches: 0  
 Query Match: 56.68% Indels: 197  
 DB: 5 Gaps: 3

US-10-631-958-11 (1-562) x AAS77730 (1-2186)

Qy 195 ThrGluHisAlaAsnGlnAlaLysGluThrLeuTrpGluLeuLeuAsnLysTyrAsp 214  
 Db 3 ACTGAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGAC 62  
 Qy 215 Gly----- 215  
 Db 63 GGATGTACTACTGCCCATTCAGGATGCCCTGTCTCGAAACCAACAGCCCGCTGGGAG 122  
 Qy 215 ----- 215  
 Db 123 GCTGCCCGGTGCAAGGTGTAGCTACGGGGAAGGCAAGAGCCTTCCCATGGAGCCG 182  
 Qy 215 ----- 215  
 Db 183 GGCCCAAGCAGATGCTGCCAAGGGCCAGGTGTGCAGAGGTGCCCTTCTTGAAGCTGGAA 242  
 Qy 215 ----- 215  
 Db 243 GCCTCAGGCTGTGCGCTCAGAGGGCCGACCTGTCAGGAGCCCTGACCGGCTGTGTTC 302  
 Qy 216 -----IleValCysValGlyAspGlyMetPheSerGluValLeuHisGlyLeu 232  
 Db 303 TCCTGCAGCATCGTCTGTGTCGGCGAGATGGTATGTTTCAGCGAGGTGCTGCACGGTCTG 362  
 Qy 233 IleGlyArgThrGlnArgSerAlaGlyValAspGlnHisProArgAlaValLeuVal 252  
 Db 363 ATTGGGAGGACGACAGAGCGCGGGGTGCAGCAACCAACCCCGGGGTGTGTCTGGTC 422  
 Qy 253 ProSerSerLeuArgIleGlyIleIleProAla----- 263  
 Db 423 CCCAGTAGCCTCCGATTTGAATCATTTCCCGCAGGCTTTCGAATGACCTGGCAGGGGAC 482  
 Qy 263 ----- 263  
 Db 483 GAGGTGTCTGTCTCTCTGCGCCTGTGTCTGCGCCCGAGGGTGGCGGCATGTTGCACACT 542  
 Qy 263 ----- 263  
 Db 543 TTCACCTCTCTCAGACTCTGGGATGTGAGACCGCAGTCATCCCATTTTATGGATGAA 602  
 Qy 264 -----GlySer 265  
 Db 603 GACAGGAGGACTGGGAGCATGTGGGCCCGGTGAGAACGCTGTGTGGCTTGGACGGGTCA 662  
 Qy 266 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 285  
 |||||||

```
Db 663 ACGGACTGCGTGTGTTACTCCACCGTGGCGACCAGGAGCAGAACCTCGGCGCTGCAT 722
Qy 286 lleValVal----- 288
Db 723 ATCGTGTGTTGGCTGTGCCCCGAGGCTCGAAACACCACCGCGTCCGCCATACCTGCTGT 782
Qy 288 ----- 288
Db 783 GCGAGTGGCGACGGCAGCTGTGCTGGCTGCCAGCTGTGGAACGCATGCCCTGTGTGAG 842
Qy 289 -----GlyAspSerLeuAlaMetAspValSer 297
Db 843 GCGTCGAGGCTTCAGTCCAGGATGCGAGGCCCGGGGACTCGCTGGCCATGATGTGCC 902
Qy 298 SerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPhe 317
Db 903 TCAGTCCACCAACAGCAGCACTCCCTCGCTACTCCGTGTCCTGCTGGGCTACGGCTTC 962
Qy 318 TyrGlyAspIleIleValAspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAsp 337
Db 963 TAGGGGACATCATCAGGACAGTGAAGAAACCGTGTGGTCTTGGCCAGATACGAC 1022
Qy 338 PheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeu 357
Db 1023 TTTTCAGGTTAAAGACCTTCCTCCACCACTGCTATGAAGGCAGTGTCTCTTCCTC 1082
Qy 358 ProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPhe 377
Db 1083 CCGTCACAACACACCGTGGGATCTCCAAGGGATAGGAAGCCCTGCCGGCAGGATGCTTT 1142
Qy 378 ValCysArgGlnSerLysGlnLeuGluGluGlnLysLysLysAlaLeuTyrGlyLeu 397
Db 1143 GTTTCGAGCAAGCAAGCAGCAGCTGGAGGAGGAGCAGAAAGCACTGTATGTGTTG 1202
Qy 398 GluAlaLysGluAspValGluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsn 417
Db 1203 GAAGTCGGGAGGACGTGGAGGATGGCAAGTCGTCTGTGGAGCTTCTGGCCATCAAT 1262
Qy 418 AlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHis 437
Db 1263 GCCAACAATATGTCCTGTGTTGCGCGGAGCCCGAGGGGCTCTCCCGCTGCCAC 1322
Qy 438 LeuGlyAspGlySerSerAspLeuIleLeuLysCysSerArgPheAsnPheLeu 457
Db 1323 TGGGAGACGGGTCTCTGACCTCATCTCCGAAATGCTCCAGGTTCATTTTCTG 1382
Qy 458 ArgPheLeuIleArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluVal 477
Db 1383 AGATTTCTCATCAGGCACACCAACACGAGCAGGACGATTTGACTTCACCTTTTGTGAAGTT 1442
Qy 478 TyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeu 497
Db 1443 TATCGCTCAAGAAATTCAGTTTACGTCGAAGCAGCATGGAGGATGAGCAGCGACCTC 1502
Qy 498 LysGluGlyGlyLysLysArgPheGlyHisLysCysSerSerHisProSerCysCys 517
Db 1503 AAGGAGGGGGGAAAGCGCTTTTGGGCACATTTGCGACAGCCACCCCTCTCGTCTGTC 1562
Qy 518 ThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGlu 537
Db 1563 ACCGTCTCCACAGCTCTCTGGNACTGCCAGCGGGGAGGTCCTGCCACAGCCCTGCCATCGAG 1622
Qy 538 ValArg 539
Db 1623 GTCAGA 1628
RESULT 14
ABL40822
ID ABL40822 standard; cDNA; 979 BP.
XX AC ABL40822;
XX DT 03-JUL-2002 (first entry)
```

```
XX Human sphingosine kinase-like protein encoding cDNA.
DE
XX
KW Human sphingosine kinase-like protein; intracellular signalling; gene;
KW cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;
KW autoimmune disease; rheumatoid arthritis; Parkinson's disease; ss.
XX
OS Homo sapiens.
XX
Key Location/Qualifiers
CDS 2..789
    /*tag= a
    /product= "sphingosine kinase-like protein"
    /note= "start and stop codons are not indicated"
XX
XX WO200228906-A2.
XX
XX 11-APR-2002.
XX
XX 05-OCT-2001; 2001WO-EP011516.
XX
XX 06-OCT-2000; 2000US-0238005P.
XX
XX 23-AUG-2001; 2001US-0314113P.
XX
XX (PARB ) BAYER AG.
XX
XX Kossida S, Encinas J;
XX
XX WPI; 2002-340094/37.
XX
XX P-PSDB; ABB07854.
XX
XX New reagent for modulating the activity of sphingosine kinase-like
XX protein polypeptide or polynucleotide and treating cancer, asthma,
XX allergy, an autoimmune disease, or a central or peripheral nervous system
XX disorder.
XX
XX Claim 1; Fig 1; 120pp; English.
XX
XX The invention relates to a human sphingosine kinase-like protein. The
XX polypeptide can be expressed by standard recombinant methodology. The
XX sphingosine kinase-like protein and gene can be used to regulate
XX intracellular signalling and consequently cell proliferation and
XX apoptosis. Such regulation is useful for treating cancer, allergies (e.g.
XX asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and
XX peripheral nervous system disorders (e.g. Parkinson's disease). The
XX present sequence represents the human sphingosine kinase-like protein
XX encoding cDNA
XX
XX Sequence 979 BP; 237 A; 259 C; 271 G; 212 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 2.32e-123 Length: 979
Score: 1640.50 Matches: 314
Percent Similarity: 96.63% Conservative: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 54.23% Indels: 11
DB: 6 Gaps: 1
XX
US-10-631-958-11 (1-562) x ABL40822 (1-979)
Qy 155 ProlYsHisLeuLeuValPheIleAsnProPheGlyLysGlyGlnGlyAsnArgile 174
Db 2 CCAAGCATTTTACTGGGTATTTATCAACCGCTTTGGAGGAAAGGACAGGCGGATA 61
Qy 175 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIle--- 193
Db 62 TATGAAAGAAAGTGGGACCACTGTTCCCTTACCTTCCATCCACTGACATCATCGGT 121
Qy 194 -----ValThrGluHisAlaAsnGlnAlaLysGlu 203
Db 122 AACAAATTCATGTTAACTATGTAGAAAGTAATTACTGAACATGCTAATCAGCCAGGAG 181
Qy 204 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 223
```

```
Db 182 ACTCTGTATGAGATTAAACATAGACAAATACACGGCATCGTCTGTCTGCGCGAGATCGT 241
Qy 224 MetPheSerGluValLeuHisGlyLeuLeuGlyValThrGlnArgSerAlaGlyValAsp 243
Db 242 ATGTTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACAGAGCGCGGGGTGCAC 301
Qy 244 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgGlyLeuGlyValLeuProAla 263
Db 302 CAGAACCAACCCCGGGCTGTCTGCTCCCAAGTAGCCTCCGGATTGGATCATTCGCCCA 361
Qy 264 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 283
Db 362 GGGTCAACCGGACTCGTGTGTACTCCACCGTGGGCACACGACGACGAGAAACCTCGCG 421
Qy 284 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 303
Db 422 CTGCATATCGTGTGTGGGACATCGCTGGCCATGATGTCTCAGTCCACCAACAGC 481
Qy 304 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 323
Db 482 ACATCTTCTGCTACTCGTGTCTCTGCTGGCTACGGCTTCTACGGGGACATCATCAAG 541
Qy 324 AspSerGluLysLysArgTyrTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 343
Db 542 GACAGTGAAGAAACCGTGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAGACC 601
Qy 344 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 363
Db 602 TTCTCTCTCCCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACAAACACACGGTG 661
Qy 364 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 383
Db 662 GGATCTCAAGGATAGGAGCCCTGCGGCGACAGATGCTTTGTTTCAGGCAAGCAAG 721
Qy 384 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 403
Db 722 CAGCAGCTGGAGGAGGACAGAAAGACATGATGTTTGGAGCTGCGCGAGGAGCTG 781
Qy 404 GluGluTrpGlnValValCysGlyValPheLeuAlaIleAlaThrAsnMetSerCys 423
Db 782 GAGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCAACCAATGCTCTGT 841
Qy 424 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 443
Db 842 GCTTGTCCCGGAGCCCGAGGGCTCTCCCGGCTGCCACTTGGGAGCGGTCTTCT 901
Qy 444 AspLeuLeuLeuIleArgLysCysSerArgPheAlaHisPheLeuArgPheLeuAlaArgHis 463
Db 902 GACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAC 961
Qy 464 ThrAsnGlnAspGln 469
Db 962 ACCAACCCAGCAGGACCAG 979
RESULT 15
ID AAS77731 standard; cDNA; 1570 BP.
XX AC AAS77731;
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #13535.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN W0200175067-A2.
XX PD 11-OCT-2001.
```

```
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG13544.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX Claim 1; SEQ ID NO 13535; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activities. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 1570 BP; 349 A; 446 C; 456 G; 319 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,65e-102 Length: 1570
Score: 1386.50 Matches: 278
Percent Similarity: 79.10% Conservative: 21
Best Local Similarity: 73.54% Mismatches: 45
Query Match: 45.83% Indels: 34
DB: 5 Gaps: 6

US-10-631-958-11 (1-562) x AAS77731 (1-1570)
Qy 186 AlaserIleThrThrAspIleIleValThrGluHisAlaAsnGlnAla-LysGluThrLe 205
Db 437 AGCAGCATCCCGCACAGCAGTGTGTGTGGACATGCCACAGCATCCGGGGAGAAAAC 496
Qy 205 uTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPh 225
Db 497 ACACCTGACCCAGGTGGACAGA-----TGCTGTGGG----- 527
Qy 225 eSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp---Gl 244
Db 528 -----AACATCTTGAAGCGCGGAGAGGCTGCTCCCCGAGGCTCG 571
Qy 244 nAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIlelleProAlaGl 264
Db 572 AAACCAACCGGTCCCGCCATACCTCT-----GTGCGAGTGGG 610
Qy 264 ySerThrAspCysValCysTyrSerThrValGlyThrSerAspAla--GluThrSerAla 283
Db 611 CACGGGAGCTGTGCTGGGCTGCCAGCTGTGAACGCATGCTCTGTGAGGCTCTCGAG 670
```



**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 6, 2005, 02:35:14 ; Search time 7402.62 Seconds  
(without alignments)  
3678.672 Million cell updates/sec

Title: US-10-631-958-11

Perfect score: 3025

Sequence: 1 HEANGPAPLGVRAAPPWRT.....QLVRLFARGIBENKPDPSHS 562

Scoring table:

BLASUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-O=/cpn2\_1/USPTO spo01/US10631958/runat\_02092005\_165815\_4533/app.query.fasta\_1.1941  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10631958 @CGN 1 1 11756 @runat\_02092005\_165815\_4533 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3025	100.0	4413	6	AX457006 Sequence
2	3012	99.6	2042	9	HS457828 Homo sapi
3	3012	99.6	4445	9	AB079066 Homo sapi
4	3004	99.3	4463	6	BD102675 Ceramide

5	2931.5	96.9	4432	6	AR541900	Sequence
6	2901	95.9	1654	9	CR456404	Homo sapi
7	2792.5	92.3	1772	9	BC067255	Homo sapi
8	2649.5	87.6	1840	6	AX224383	Sequence
9	2588	85.6	4171	6	BD183468	Novel gen
10	2588	85.6	4171	9	AB051433	Homo sapi
11	2510	83.0	1459	6	CQ730476	Sequence
12	2436.5	80.5	2830	10	AB079067	Mus muscu
13	1912.5	63.2	3661	10	AK129416	Gallus ga
14	1798	59.4	1450	5	CR386590	Sequence
15	1714	56.7	2494	5	BC074350	Xenopus l
16	1640.5	54.2	979	6	AX456998	Sequence
17	1496	49.5	1520	5	BC074110	Xenopus l
18	1008	33.3	550	6	AX457005	Sequence
19	886.5	29.3	2333	3	AK112750	Ciona int
20	813	26.9	2064	6	AR509346	Sequence
21	813	26.9	3178	3	BT014925	Drosophil
22	811	26.8	474	6	AX457001	Sequence
23	763	25.2	1810	3	AY061001	Drosophil
24	721.5	23.9	10337	6	CQ608039	Sequence
25	721.5	23.9	22353	2	AC017269	Drosophil
26	721.5	23.9	181009	3	AC007575	Drosophil
27	721.5	23.9	295289	3	AE003603	Drosophil
28	686.5	22.7	2173	6	CQ608040	Sequence
29	656	21.7	522	6	AX224385	Sequence
30	648	21.4	2257	8	AY362552	Arabidops
31	622.5	20.6	2263	8	AY035126	Arabidops
32	595	19.7	1599	9	AY357073	Homo sapi
33	585.5	19.4	1084	6	AR510195	Sequence
34	543	18.0	3039	9	BC004278	Homo sapi
35	543	18.0	3074	9	HS457002	Homo sapi
36	541.5	17.9	70890	9	HS457002	Human DNA
37	525	17.4	329	6	AX457002	Sequence
38	511.5	16.9	181772	2	AC007689	Homo sapi
39	498	16.5	260331	2	AC135400	Rattus no
40	450	14.9	240483	2	AC095955	Rattus no
41	437.5	14.5	564	6	AX718996	Sequence
42	430	14.2	1799	8	AK072068	Oryza sat
43	430	14.2	1817	8	AK107062	Oryza sat
44	428	14.1	1458	8	BT006621	Arabidops
45	428	14.1	1757	8	AY128394	Arabidops

ALIGNMENTS

RESULT 1	AX457006	Sequence 9 from Patent WO0228906.	4413 bp	DNA	linear	PAT 06-JUL-2002
LOCUS	AX457006	AX457006	4413 bp	DNA	linear	PAT 06-JUL-2002
DEFINITION	AX457006	AX457006	4413 bp	DNA	linear	PAT 06-JUL-2002
ACCESSION	AX457006	AX457006	4413 bp	DNA	linear	PAT 06-JUL-2002
VERSION	AX457006.1	GI:21715795				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	1					
AUTHORS	Kossida, S. and Encinas, J.					
TITLE	Regulation of human sphingosine kinase-like protein					
JOURNAL	Patent: WO 0228906-A 9 11-APR-2002;					
FEATURES	Bayer Aktiengesellschaft (DE)					
source	Location/Qualifiers					
	1. 4413					
	/organism="Homo sapiens"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:9606"					
	76. 1689					
	/note="unnamed protein product"					
	/codon_start=1					
	/protein_id="CAD38256.1"					
	/db_xref="GI:21715796"					
	/translation="MGATGAPEPLQSLVWVQKRCANVLEPARALLRWRSFGPGAGA					
	PGADACSPVPVSEIIAVBETDVHGKQSGKQWMEKPYAFTVHCVKRRHRWKAQV					

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

CDS

TFWCPBQLCHLWLTQIREMLEKLTSRPKHLVFINPFGKGQKRIYERKVAFLPTL  
ASITFTDIIVTEHANAQKETIYEINIDYDGIIVCVGDMFSEVLHGLIGRTORSAGVD  
QHPRVAVLSSLRIGIIPAGSTDCVSVTGTSDAETSALHIWVGSLAMDVSSVHH  
NSTLLRYSLGLGFGVDIIKDKSEKKRWGLGARYDPSGLKTFLSHHCYGTVSFLPA  
QHTYSPRDRKPCRCAGCFVCRQSQJLEBEOKKALYGLEAAEDVEEHQVCGKFLAIN  
QTNMSACRRSPRGLSPAHLHGDSGLIIRKCRNFNLFRLIRHNNQDDQDFTV  
EYVRVKFQTSKMEDESDLKEGGKKRFGHICSSHPSCCCTVSNSWNCDEVLHS  
PAIEVRVHCQLVRLFARIBENPKPDHS\*

ORIGIN

Alignment Scores:  
Pred. No.: 3,33e-199 Length: 4413  
Score: 3025.00 Matches: 562  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-631-958-11 (1-562) x AX457006 (1-4413)

Qy	1	HisGluAlaAAsnGlyProAlaProLeuGlyValArgAlaProProAlaTrpArgThr	20
Db	1	CACAGGCGGTAAAGTCCGGCGCCCTCGCGCTCCGGCGCCCGCCAGCGCTGGCGAGC	60
Qy	21	SerProAlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrp	40
Db	61	AGCCCGCGCGAGATGGGGCGCAGCGGGCGCGGAGCGGCTGCAATCGTCTGTGG	120
Qy	41	VallysGlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrp	60
Db	121	GTGAAGCAGCAGCGCTGCGCGCTGAGCTGAGCGCCGCGCGGCTCTGCTGCGCTGTGG	180
Qy	61	ArgSerProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSer	80
Db	181	CGGAGCCCGGGCCGCGAGCCGCGCCCGCGCGGATGCTCTGTGCGCTGTATCT	240
Qy	81	GluIleLeAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrp	100
Db	241	GAGATCATCGCGTTGAGGAAACAGACGTTCAACGGGAAACATCAAGCAGTGGGAAATGG	300
Qy	101	GlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgArgHis	120
Db	301	CAGAAATGGGAAAGCCCTTACGCTTTACAGCTTCACTGTGTAAAGAGACGACGCGCAC	360
Qy	121	ArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrp	140
Db	361	CGCTGGAAAGTGGGCGCAGTGACCTTTCTGGTGTCCAGGAGCAGCTGTCTACTGTGG	420
Qy	141	LeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuVal	160
Db	421	CTGCAGACCCCTGCGGAGATCTGGAGAAAGCTGACGCTCCAGACCAAGCATTTACTGGTA	480
Qy	161	PheIleAsnProPheGlyGlyGlyGlnGlyLysArgIleTyrGluArgLysValAla	180
Db	481	TTTATCAACCCGTTTGGAGGAAAGGACAAAGCGGATATATGAAAGAAAGTGGCA	540
Qy	181	ProLeuPheThrLeuAlaSerIleThrThrAspIleLeuValThrGluHisAlaAsnGln	200
Db	541	CCACTGTTTCACTTAGCCCTCCATACCACTGACATCATCGTTACTGAAACATGCTTAATCAG	600
Qy	201	AlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGly	220
Db	601	GCCAAGGAGACTCTGTATGAGATTAAACATACAAAATACGACGCGCATCTGTGTGCGC	660
Qy	221	GlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAla	240
Db	661	GGAGATGGTATGTTCAGCAGGTGCTGCACGGTCTGATGGGAGGACGCGAGAGCGCC	720
Qy	241	GlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIle	260
Db	721	GGGGTCGACCAAGAACCCCGGGCTGTGTGGTCCCGAGTAGCCCTCGGATGGGAATC	780
Qy	261	IleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGlu	280

Db	781	ATTCCCGCAGGGTCAACGGACTCGTGTGTGTACTCCACCGTGGGCACACGCGCAGAA	840
Qy	281	ThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerValHis	300
Db	841	ACCTCGCGCTGCATATCGTGTGTGGGACTCCGTGGCCATGGATGTGTCTCAGTCCAC	900
Qy	301	HisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAsp	320
Db	901	CACACAGCACACTCCCTTCGTACTCCGTGTCTCTGCTGGGCTACGGCTTCTACGGGAC	960
Qy	321	IleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGly	340
Db	961	ATCATCAAGGACAGTGAAGAAACGGTGGTGGTCTGGTCTGGCCAGATACGACTTTTCAGGT	1020
Qy	341	LeuLysThrPheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAlaGln	360
Db	1021	TTAAAGACCTTCTCTCCCACTCTATGAAGGACAGTGTCTCTCTCTCCCTGCACAA	1080
Qy	361	HisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArg	380
Db	1081	CACAGGTGGGATCTCCAGGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTGCAGG	1140
Qy	381	GlnSerLysGlnGlnLeuGluGluGlnLysAlaLeuTyrGlyLeuGluAlaAla	400
Db	1141	CAAGCAAGCAGCAGCTGGAGGAGGACAGAGAAACACTGTATGTGTTTGGAGGCTGCG	1200
Qy	401	GluAspValGluGluTrpGlnValValCysGlyValPheLeuAlaIleAsnAlaThrAsn	420
Db	1201	GAGACGTGGAGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCAANAAC	1260
Qy	421	MetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAsp	440
Db	1261	ATGTCCTGTCTTGTCCCGGAGCCCGAGGGGCTCTCCCCGGCTGCCCCCTTTGGGAGAC	1320
Qy	441	GlySerSerAspLeuIleLeuArgLysCysSerArgPheAsnPheLeuArgPheLeu	460
Db	1321	GGGTCTTCTGACCTCATCTCATCCGGAATGTCTCCAGGTTCATTTTCTGAGATTTCTC	1380
Qy	461	IleArgHisThrAsnGlnAspGlnPheAspPheThrPheValGluValTyrArgVal	480
Db	1381	ATCAGGCACACCAACACGAGGACCACTTTGACTTCACTTTTGTGAAGTTTATCGCGTC	1440
Qy	481	LysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGly	500
Db	1441	AAGAAATTCAGTTTACGTGGAAGCACATGAGGATGAGGACAGCGACCTCAAGAGGGG	1500
Qy	501	GlyLysLysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSer	520
Db	1501	GGGAAGAAGCGCTTGGGCACATTTGCAGCAGCACCCCTCTCTGCTGCTGACCGTCTCC	1560
Qy	521	AsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgVal	540
Db	1561	AAAGCTCTGGAACTCGCAGCGGGAGGTCTGCACAGCCCTGCCATCGAGGTCAGAGTC	1620
Qy	541	HisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSer	560
Db	1621	CACGCGAGCTGTGCTGACTCTTTTGACGAGGAATTGAAGAGAAATCCGAAGCCAGACTCA	1680
Qy	561	HisSer 562	
Db	1681	CACAGC 1686	

RESULT 2

HSA457828 2042 bp mRNA linear PRI 19-APR-2002  
LOCUS HSA457828  
DEFINITION Homo sapiens mRNA for putative lipid kinase (LK4 gene).  
ACCESSION AJ457828  
VERSION AJ457828.1 GI:20269072  
KEYWORDS lipid kinase; LK4 gene.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Van Veldhoven, P.P.  
A search for lipid kinases  
Unpublished  
2 (bases 1 to 2042)  
Van Veldhoven, P.P.  
Direct Submission  
Submitted (18-APR-2002) Van Veldhoven P.P., Farmakologie,  
K.U.Leuven, Herestraat, B-3000 Leuven, BELGIUM  
FEATURES  
source  
1..2042  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="22"  
/map="22q13"  
1..2042  
/gene="LK4"  
71..1684  
/gene="LK4"  
/codon\_start=1  
/product="putative lipid kinase"  
/protein\_id="CAD29884.1"  
/db\_xref="GI:20269073"  
/db\_xref="GOA:Q8TCT0"  
/translation="MGATCAAEPLQSLVWVQQRCAVSLSEPARALLRWRSPGPCAGA  
PGADACSVPSVEI IAVETDVHGKQSGKQKMEKPYFTVCHVCRRARRRKKWAQV  
TFWCPEROLCHLWLOTLREMLEKLTSPKHLVFINPFGKQGRKIVYERKVPFTL  
ASITDIIIVTEHNAQAKETLYEINIDKIDGIVCVGGDNFSEVLHGLIGRTQRSAGVD  
QNHPRVLPSSLRIGIIPAGSTDCYSTVSTGSDAETSALHI VVGDLSLMDVSSVHH  
NSTLARYSVLLGYFGDITKSEKKRWGLARYDFSLKTFILSHHCYEGTVSFLPA  
QHTVGSPRDRKPRACGFCVQSKQLEEKQKALYGLEAAEDVEEVMQVCGKFLAIN  
ATNMSCACRRSPRGLSPAHLGDGSSDLILRKCSRNFRLRLRHTNQDQDFTFV  
EYVRVKQFQTSKHWEDSDSLKEGKKRFGHICSSHPSCCCTVSNSSNWCNCGEVLHS  
PAIEVRVHCQLVRLFARGIEENPKFDSHS"

ORIGIN

Alignment Scores:  
Pred. No.: 1e-198 Length: 2042  
Score: 3012.00 Matches: 560  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.57% Indels: 0  
DB: Gaps: 0

US-10-631-958-11 (1-562) x HSA457828 (1-2042)

QY 3 AlaAlaAsnGlyProAlaProLeuGlyValArgAlaProProAlaTyrArgThrSerPro 22  
DB 2 GCGGCTAACGGTCCGGCGCCCTCGCGCTCCGGCGCCGCCAGCCTGGCGGACGACCG 61  
QY 23 AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTyrVallys 42  
DB 62 GCGGCGGAGATGGGGCGCGGGCGGGCGGGAGCCGCTGCAATCCGCTGTGTGGGTGAAG 121  
QY 43 GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTyrTyrArgSer 62  
DB 122 CAGCAGCGCTGCGCGTGAAGCTTGAGCCCGCGCGCGGGCTCTGCTGGCTGGTGGCGGAGC 181  
QY 63 ProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIle 82  
DB 182 CCGGGCGCGGAGCGCGCGCCCGCGCGGATGCTGCTGTGCTGTATCTGAGATC 241  
QY 83 IleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTyrGlnLys 102  
DB 242 ATCGCGGTGGAGAAACACAGCTTCACGGGAAACATCAAGGCAGTGGAAAAATGGCAGAAA 301  
QY 103 MetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTyr 122  
DB 302 ATGGAAAGACCTTACGCTTTTACAGTTTCACTGTGTGTATAGAGAGACGACGCGCACCGCTGG 361

QY 123 LysTyrAlaGlnValThrPheTyrCysProGluGlnLeuCysHisLeuTyrLeuGln 142  
DB 362 AAGTGGCGCGCAGGTGACTTTCTGTGTGCCAGAGAGCAGCTGTGTCTACTTGTGGTTCGAG 421  
QY 143 ThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIle 162  
DB 422 ACCCTGCGGAGATGCTGGAGAGCTGACGCTCCAGACCAAGACATTTACTGGTATTATC 481  
QY 163 AsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeu 182  
DB 482 AACCCGTTTGGAGAAAGGACCAAGCGGATATATGAAGAAAAAGTGGCACCACCTG 541  
QY 183 PheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLys 202  
DB 542 TTCACCTTAGCTCCATCACCACATGACATCATCTGTACTTGAACATGCTTAATCAGGCCAAG 601  
QY 203 GluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyValAsp 222  
DB 602 GAGACTCTGTATGAGATTACATAGACAAATACACCGGCATCGTCTGTGTGGCGGAGAT 661  
QY 223 GlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyVal 242  
DB 662 GGTATGTTACGAGGTGCTGCACGGTCTGATTGGAGGACCGACAGAGAGCGCCGGGTC 721  
QY 243 AspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIlePro 262  
DB 722 GACCAGAACCAACCCCGGGCTGTGCTGCTCCCATGAGCTCCGGATTTGGAAATCATTC 781  
QY 263 AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer 282  
DB 782 GCAGGCTCAACGACTGCGTGTGTATTCTCCACCGTGGGCACGAGGAGCGAGAACCTCG 841  
QY 283 AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsn 302  
DB 842 CGCTGCATATCGTTGTTGGGACTCGCTGGCCCATGATGTGTCTCAGTCCACCAAC 901  
QY 303 SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIle 322  
DB 902 AGCACATCT 961  
QY 323 LysAspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys 342  
DB 962 AAGGACAGTGAAGAAGAAACGGTGTGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAG 1021  
QY 343 ThrPheLeuSerHisCysTyrGlyThrValSerPheLeuProAlaGlnHisThr 362  
DB 1022 ACCTTCTCTCCACCCTGCTATGAAGGACAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1081  
QY 363 ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer 382  
DB 1082 GTGGGATCTCCAAAGGATAGGAAGCCCTGCGCGGCGAGGATGCTTTGTTTGGAGGCAAGC 1141  
QY 383 LysGlnGlnLeuGluGlnLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp 402  
DB 1142 AAGCAGCAGCTGGAGGAGGAGCAGAGAACACTGTATGTGTTTGGAGCTGCGGAGGAC 1201  
QY 403 ValGluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSer 422  
DB 1202 GTGGGAGGTGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCC 1261  
QY 423 CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer 442  
DB 1262 TGTGCTTGTGCGCGGAGCCCGAGGGGCTCTCTCCCGGCTGCGCCACTTGGGAGAGCGGGTCT 1321  
QY 443 SerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArg 462  
DB 1322 TCTGACCTCATCTCATCCGGAATGTCTCAGGTTCAATTTTCTGAGATTCTCATCAGG 1381  
QY 463 HisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLys 482  
DB 1382 CACACCAACACGAGGACGAGTCTTGTGACTTCTCACTTTTGTGAAGTTTATCGTCGTCAGAAA 1441  
QY 483 PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLys 502



Db	1075																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
----	------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

source		1. .4463	/organism="Homo sapiens"
			/mol_type="genomic DNA"
			/db_xref="taxon:9606"
ORIGIN			
Alignment Scores:			
Pred. No.:	9.55e-198	Length:	4463
Score:	3004.00	Matches:	558
Percent Similarity:	99.82%	Conservative:	1
Best Local Similarity:	99.64%	Mismatches:	1
Query Match:	99.31%	Indels:	0
DB:	6	Gaps:	0
US-10-631-958-11 (1-562) x BD102675 (1-4463)			
Qy	3	AlaAlaAsnGlyProAlaProLeuGlyValArgAlaProProAlaTrpArgThrSerPro	22
Db	55	GCGCTAAAGGTCCGGGGCCCTCGGGCTCGGGGCCCGCCCGAGCTGGCGACGAGCCCG	114
Qy	23	AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLys	42
Db	115	GCGCGGAGATGGGGGCGACGGGGCGCGGAGCGCTGCAATCGTCTGTGGGTGAAG	174
Qy	43	GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSer	62
Db	175	CAGCAGCGCTGCGCGGTGAGCTTGGAGCCCGCGGGCTCTGCTGCGCTGGTGGCGGAGC	234
Qy	63	ProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIle	82
Db	235	CCGGGCCCCGAGCGGGGCCCGCCCGCGGATGCTGCTCTGTGCTGTATCTGAGATC	294
Qy	83	IleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLys	102
Db	295	ATCGCGCTTGAGGAAACAGACGTTTACGGGAAAATCAAGGAGTGGGAAAATGGCAGAA	354
Qy	103	MetGluLysProTyraAlaPheThrValHisCysValLysArgAlaArgHisArgTrp	122
Db	355	ATGAAAAGCGTTACGCTTTTACGTTTCTGTTTAAAGAGACGACGACCGCTGG	414
Qy	123	LysTrpAlaGlnValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGln	142
Db	415	AACTGGGCGAGGTGACTTTCTGGTGTCCAGAGGAGGAGCTGTGTCTACTTGTGCTGAG	474
Qy	143	ThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIle	162
Db	475	ACCTCGCGGAGATGCTGGAGAACTGAGCTCCAGACCAAGCATTTACTGGTATTTATC	534
Qy	163	AsnProPheGlyLysGlyGlnGlyLysArgIleTyrgluArgLysValAlaProLeu	182
Db	535	AACCCGTTTGGAGGAAAAGGCAAGGCAAGCGGATATATGAAGAGAAAGTGGCACCTG	594
Qy	183	PheThrLeuAlaSerIleThrThrAspIleValThrGluHisAlaAsnGlnAlaLys	202
Db	595	TTTACCTTAGCTCCATCACCACCTGACATCATCGTTTACTGAACATGCTTAATCAGGCCAAG	654
Qy	203	GluThrLeuTyrgluIleAsnIleAspLysTyrglyIleValCysValGlyGlyAsp	222
Db	655	GAGACTCTGATGAGATTAAACATAGACAAATACACGCGCATCGTCTGTGTCGGGGAGAT	714
Qy	223	GlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyVal	242
Db	715	GGTATGTTTACGAGGTGCTGCGCGTGTGTTTGGGAGGACGAGAGAGCGCGGGGTC	774
Qy	243	AspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIlePro	262
Db	775	GACCAGAACCAACCCCGGGCTGTGCTGCTCCAGTAGCTCCGGATTGGAAATCATTTCCC	834
Qy	263	AlaGlySerThrAspCysValCysTyrsrThrValGlyThrSerAspAlaGluThrSer	282
Db	835	GCAGGTCCACGGACTGCGTGTGTTACTCCACCGTGGGACACGAGCGCAGAAACCTCG	894
Qy	283	AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsn	302

Db	895	GCCTGTCATATCGTTGTTGGGACATCGCTGCGCCATGATGTCTCAGTCCACCAAC	954	
Qy	303	SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIle	322	
Db	955	AGCACATCTCTTCGTACTCGCTGCTCCCTGCTGGGCTACGGCTTCTACGGGACATCATC	1014	
Qy	323	LysAspSerGluLysValArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys	342	
Db	1015	AAGACAGTGAAGAAGAAACGGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAG	1074	
Qy	343	ThrPheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr	362	
Db	1075	ACCTTCTCTCCACCACTGCTATGAAGGACAGTCTCTTCTCTCCCTGCACACACACG	1134	
Qy	363	ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer	382	
Db	1135	GTGGGATCTCCAAAGGATAGGAAGCCCTGCGCGCAGGATGCTTTGTTGAGCAAGC	1194	
Qy	383	LysGlnGlnLeuGluGluGlnLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp	402	
Db	1195	AAGCAGACGCTGGAGGAGCAGAGAAACACTGTATGGTTTGGAAAGCTGCGGAGAC	1254	
Qy	403	ValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSer	422	
Db	1255	GTGGAGGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCC	1314	
Qy	423	CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer	442	
Db	1315	TGTGCTTGTGCGCGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCT	1374	
Qy	443	SerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArg	462	
Db	1375	TCTGACCTCATCTCATCCGAAGATGCTCCAAAGTTCAATTTTCTGAGATTTCTCATCAGG	1434	
Qy	463	HisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLys	482	
Db	1435	CACACCAACAGCAGGACCACTTTGACTTCACTTTTGTGAAGTTTATCGCGTCAAGAA	1494	
Qy	483	PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLys	502	
Db	1495	TTCCAGTTTACGTCCGAAGCACAATGAGGATGAGGACAGCAGCCTCAAGAGGGGGGAAG	1554	
Qy	503	LysArgPheGlyHisIleCysSerSerHisProSerCysCysCysThrValSerAsnSer	522	
Db	1555	AAGGCTTTGGGCACATTTGGCAGCAGCCACCCCTCTGCTGCTGCACCGTCTCCAAACAGC	1614	
Qy	523	SerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCys	542	
Db	1615	TCCTTGGAACTCGCATGGGAGGTCTGTCACAGCCCTGCGCATCGAGGTCAGAGTCCACTGC	1674	
Qy	543	GlnLeuValArgLeuPheAlaArgGlyIleGluGlnAsnProLysProAspSerHisSer	562	
Db	1675	CAGCTGGTTGCACTCTTTGCAAGGAATTTGAAGAGATCCGAAGCCAGACTCACACAGC	1734	
RESULT 5				
AR541900				
LOCUS	AR541900	4432 bp	DNA	linear
DEFINITION	Sequence 148 from patent US 6743619.			
ACCESSION	AR541900			
VERSION	AR541900.1	GI:53933980		
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	1 (bases 1 to 4432)			
AUTHORS	Tang, Y. T., Zhou, P., Goodrich, R., Liu, C., Asundi, V., Ren, F., Zhang, J., Zhao, Q. A., Yang, Y., Xue, A. J., Wehrman, T., Wang, J. -R., Wang, D. and Drmanac, R. T.			
TITLE	Nucleic acids and polypeptides			
JOURNAL	Patent: US 6743619-A 148 01-JUN-2004;			
FEATURES	Location/Qualifiers			
	1..4432			
	source			

/organism="unknown"					
/mol_type="genomic DNA"					
ORIGIN					
Alignment Scores:					
Pred. No.:	9.66e-193	Length:	4432		
Score:	2931.50	Matches:	548		
Percent Similarity:	98.21%	Conservative:	2		
Best Local Similarity:	97.86%	Mismatches:	7		
Query Match:	96.91%	Indels:	3		
DB:	6	Gaps:	1		
US-10-631-958-11 (1-562) x AR541900 (1-4432)					
Qy	6	GlyProAlaProLeuGly	-----ValArgAlaProProAlaTrpArgThrSerPro	22	
Db	1135	GGTTCGCGCCTATAGACAAAGACAGTGCACAGGAGCTGGCAGGCTCGCGACGAGCCCG	1194		
Qy	23	AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLys	42		
Db	1195	GCGCGGAGATGGGGCGACGCGGGCGGCGGAGCCGCTGCAATCCGTGCTGTGGGTGAAG	1254		
Qy	43	GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSer	62		
Db	1255	CAGCAGCGCTGCGCCGTGAGCTGGAGCCCGCGGGCTCTGCTGCGCTGTGGCGGAGC	1314		
Qy	63	ProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIle	82		
Db	1315	CCGGGGCCGAGCCGCGGCCCGCGGCGGATGCTGCTCTGTGCTGTATCTGAGATC	1374		
Qy	83	IleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLys	102		
Db	1375	ATCGCGCTTGAGGAAACAGACGTTTCACGGGAAACATCAAGGCGAGTGGAAAAATGGCAGAAA	1434		
Qy	103	MetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrp	122		
Db	1435	ATGGAAGAGCTTTACGCTTTTACAGTTTCACTGTGTAAAGAGAGCAGCAGCGCACCCCTGG	1494		
Qy	123	LysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGln	142		
Db	1495	AAGTGGGCGAGGTGACTTTCGTGTCCAGAGGAGCAGCTGTGTCACTTGTGGCTGCAG	1554		
Qy	143	ThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIle	162		
Db	1555	ACCTGCGGAGATGCTGGAGAGCTGACGTCCAGACCAAGCATTTTACTGGTATTATTATC	1614		
Qy	163	AsnProPheGlyLysGlyGlnGlyArgIleTyrGluArgLysValAlaProLeu	182		
Db	1615	AACCGGTTTGAGGAAAGGACAAAGGCAAGCGGATATATGAAGAAGAAAGTGGCACCACTG	1674		
Qy	183	PheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLys	202		
Db	1675	TTCACTTACCTCCATCCACTGACATCATCTGTTACTGAAACATGCTTAATCAAGGCCAAG	1734		
Qy	203	GluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyLysAsp	222		
Db	1735	GAGACTCTGTATGAGATTAAACATAGACAAATACACGGCATCGTCTGTGTCGGCGGAGAT	1794		
Qy	223	GlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyVal	242		
Db	1795	GGTATGTTTCAGCGAGGTGCTGCAAGGCTGATTGGGAGGACGACAGAGGAGCGCGGGGTC	1854		
Qy	243	AspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIlePro	262		
Db	1855	GACCAGAACCAACCCCGGCTGTGCTGCTCCCAAGTAGCTCCCGATTGGAAATCATTTCCC	1914		
Qy	263	AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer	282		
Db	1915	GCAGGGTCAACGGACTGCGTGTGTACTCCACCTGGGCAACGAGCGCAGAGAAACCTCG	1974		
Qy	283	AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsn	302		
Db	1975	GCGCTGCATATCGTTGTGGGAGCTCGCTGGCCATGGATGTGTCTCTCAGTCCACCAAC	2034		



Qy 303 SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIle 322  
Db 2035 AGCACACCTTCCTCGCTACTCCGTGTCTGCTGGCTACGGCTTCTACGGGACATCATC 2094  
Qy 323 LysAspSerGluLeuValArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys 342  
Db 2095 AAGGACATGAGAGAAAGCGTGGTCTTGCCAGATACGACTTTCAGGTTCAGGTTTAAAG 2154  
Qy 343 ThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr 362  
Db 2155 ACCTTCCTCTCCACCACTGCTATGAGGAGCAGTGTCTTCTCTCCCTGCACACACACAG 2214  
Qy 363 ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer 382  
Db 2215 GTGGGATCTCAAGGAGATAGGAAGCCCTGCCGGCAGGATGCTTGTGTTCAGGCAAGC 2274  
Qy 383 LysGlnGlnLeuGluGluGlnLysValAlaLeuTyrGlyLeuGluAlaAlaGluAsp 402  
Db 2275 AAGCAGCAGCTGGAGAGAGCAGAGAAAGCAGCTGTATGTTTGGAAAGCTGGCAGGAC 2334  
Qy 403 ValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSer 422  
Db 2335 GTGGAGGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCC 2394  
Qy 423 CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer 442  
Db 2395 TGTGCTTGTCCCGGAGCCCGAGGGCCCTCTCCCGGCTGCCACTTGGGAGAGGGGTCT 2454  
Qy 443 SerAspLeuLeuLeuLeuArgLysCysSerArgPheAenPheLeuArgPheLeuLeuArg 462  
Db 2455 TCTGACCTCATCTCTCATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGG 2514  
Qy 463 HisThrAenGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLys 482  
Db 2515 CACACCAACAGCAGGACCACTTGTGACCTTCTACTTTTGTGAAGTTTATCGGTCAAGAA 2574  
Qy 483 PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyGlyLys 502  
Db 2575 TTCCAGTTTACGTTCGAAGCACATGAGGATGAGGACAGCAGCTCAAGAGGGGGGAAG 2634  
Qy 503 LysArgPheGlyHisIleCysSerSerHisProSerCysCysCysThrValSerAsnSer 522  
Db 2635 AAGCGCTTTGGGCACATTTGACAGCAGCCACCCCTCTCTGCTGTCACCGTCTCAACAGC 2694  
Qy 523 SerTrpAenCysAspGlyValLeuHisSerProAlaIleGluValArgValHisCys 542  
Db 2695 TCCTGGAACCTGGACGGGAGGTCTGCACAGCCCTGCCATCGAGTCCAGTCCACTGC 2754  
Qy 543 GlnLeuValArgLeuPheAlaArgGlyIleGluGluAenProLysProAspSerHisSer 562  
Db 2755 CAGCTGTTTCGACTCTTTTGCACGAGGATTCGAAGAGATCCGAAGCCAGACTCACACAGC 2814

RESULT 6  
CR456404  
LOCUS CR456404 1654 bp mRNA linear PRI 25-MAY-2004  
DEFINITION Homo sapiens bk29f11.1 full length open reading frame (ORF) cDNA  
clone (cDNA clone C22ORF.pGEM.bk29f11.1).  
ACCESSION CR456404  
VERSION CR456404.1 GI:47678338  
KEYWORDS CDNA; chromosome 22; ORF.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1654)  
AUTHORS Collins,J.E., Wright,C.L., Edwards,C.A., Davis,M.P., Grinham,J.A.,  
Cole,C.G., Goward,M.E., Aguado,B., Mallia,M., Mokrab,Y.,  
Huckle,E.J., Beare,D.M. and Dunham,I.  
TITLE Direct Submission  
JOURNAL Submitted (24-MAY-2004) Sanger Institute, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: c22@sanger.ac.uk Manuscript  
COMMENT Sanger Institute name : pGEM.bk29f11.1

Homo sapiens cDNA sequence. This sequence was generated as part of  
The Wellcome Trust Sanger Institute program to isolate cDNA clones  
representing the full length open reading frame of well annotated  
protein coding genes on human chromosome 22. For more information  
see <http://www.sanger.ac.uk/HGP/Chr22/>.

FEATURES  
source  
1..1654  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="22"  
/clone="pGEM.bk29f11.1"  
/lab\_host="JM109"  
1..1654  
/gene="bk29f11.1"  
11..1624  
/gene="bk29f11.1"  
/codon\_start=1  
/protein\_id="CAG30290.1"  
/db\_xref="GI:47678338"  
/translation="MGATGAAPLQSVLVWVQQRCAVLSLEPARALLLRWRSPPGAGAG  
PGADACSPVSEIIAVEETDVHGKHQSGWKQMEKPYAFTHVCKGARRHRWKAQV  
TFWPERQLCHLQTLERLEMLKTSRPHLLVFINFGKGQCKRIVKRVAPLFTL  
ASITDIIIVTEHANOAKETLYEINIDKVDGIVCVGGDMFSEVLHGLIGRTQSGAVD  
QNPRAVLVPSSLRIGIIPAGSTDCVCTVFTGSDAETSAHIVVDSLANDVSSVHH  
NSTLLRYSVSLGTGFYGDIIKQSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFLPA  
QHTVSGPRDKPCRCAGFCVQSKQLEBEQKALYGLAEADVQVWVCGKFLAIN  
ATNNSCARSPRGLSPAHLGGSSDLILIRKCSRFRFLRLHRTNQDQDFTFV  
EVTVKKFOFTSKHMEDESDLKEGGKKRFGHICSSHPSCCTTVSNSSNWCNDCGEVLHS  
PAIEVRVHCQLVRLPARGIENPKPDSHS"

ORIGIN  
Alignment Scores:  
Pred. No.: 3.59e-191 Length: 1654  
Score: 2901.00 Matches: 540  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 95.90% Indels: 0  
DB: Gaps: 0  
US-10-631-958-11 (1-562) x CR456404 (1-1654)  
Qy 23 AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLys 42  
Db 2 CGCGCGAGATGGGGGCGACGGGGCGCGAGCCGTCGCAATCCGTCGCTGGGTGAAG 61  
Qy 43 GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTyrTrpArgSer 62  
Db 62 CAGCAGCGCTCGCCGCTGAGCTGGAGCCCGCGGGCTCTGCTGCGCTGGTGGGAGC 121  
Qy 63 ProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIle 82  
Db 122 CGGGGGCCCGAGCGCGGCCCGCCGCGGATGCTGCTGCTGCTATCTGAGATC 181  
Qy 83 IleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLys 102  
Db 182 ATCCCGCTTGAGGAAACAGACGTTTACGGGAAACATCAAGCGCAGTGGAAAAATGCGAGAA 241  
Qy 103 MetGluLysProTyrAlaPheThrValHisCysValLysArgAlaAtgAtgHisArgTrp 122  
Db 242 ATGAAAAGCTTACGCTTTTACAGTTTACCTGTGTAAGAGAGACGACGACCGCTGG 301  
Qy 123 LysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGln 142  
Db 302 AAGTGGCGCAGGTGACTTTCTGGTGTCCAGAGGAGCAGCTGTCTCACTTGTGGCTGAG 361  
Qy 143 ThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIle 162  
Db 362 ACCCTCGGGAGATGCTGGAGAAGCTACGCTCCAGACCAAGCACTTACTGGTATTATC 421  
Qy 163 AsnProPheGlyGlyLysGlnGlyLysArgIleTyrGluArgLysValAlaProLeu 182  
Db 422 AACCCGTTTGAGGAAAAAGGACAAAGCGGATATATGAAGAAAAAGTGGCCACTG 481



Qy	183	PheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAenGlnAlaLys	202
Db	482	TTCACTTAGCTCCATCACCACCTGACATCATCGTTACTGAACATGCTAATCAGGCCAAG	541
Qy	203	GluThrLeuTyrGluIleAenIleAspLysTyrAspGlyIleValCysValGlyAsp	222
Db	542	GAGACTCTGATGAGATTAAACATAGACAAATACAGCGCATCGTCTGTCTGCGGGAGAT	601
Qy	223	GlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyVal	242
Db	602	GGTATGTTACGAGGTCCTGCACGGTCTGATTGGAGGACGACAGAGCGCGGGGTC	661
Qy	243	AspGlnAenHisProArgAlaValLeuValProSerSerLeuArgIleGlyIlePro	262
Db	662	GACCAGAACCCCGCGGCTGCTGCTCCCCAGTAGCCTCGGATTGGAATCATTTCCC	721
Qy	263	AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer	282
Db	722	GCAGGGTCAACGGACTCGTGTGTACTCCACCGTGGGCACACGCGACGAGAAACCTCG	781
Qy	283	AlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsn	302
Db	782	CGCTGTCATATCGTGTGGGACCTCGCTGGCCATGGATGTGCTCAGTCCACCAAC	841
Qy	303	SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIle	322
Db	842	AGCACACTCTTCGCTACTCCGTGCTCCTGCTGGCTACGGCTTCTACGGGGACATCATC	901
Qy	323	LysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys	342
Db	902	AAGGACAGTGAAGAAACGCTGCTGGTCTTCCAGATACGACTTTTCAGGTTTAAAG	961
Qy	343	ThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr	362
Db	962	ACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACACACACG	1021
Qy	363	ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer	382
Db	1022	GTGGGATCTCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTCAGGCAAAGC	1081
Qy	383	LysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp	402
Db	1082	AAGCAGCAGCTGGAGGAGGACAGAGAAGACACTGTATGTTTGGAAAGCTGCGGAGAC	1141
Qy	403	ValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAenAlaThrAenMetSer	422
Db	1142	GTGGAGAGTGCCAAAGTCGTCGTGGGAAGTTTCTGGCCATCAATGCCAACAATGTCC	1201
Qy	423	CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer	442
Db	1202	TGTGCTTGTCCCGAGCCCAAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCT	1261
Qy	443	SerAspLeuIleLeuIleArgLysCysSerArgPheAenPheLeuArgPheLeuIleArg	462
Db	1262	TCTGACCTCATCTCATCCGAAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGG	1321
Qy	463	HisThrAenGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLys	482
Db	1322	CACACCAACGACGAGGACCACTTGCATCTCTTTGTTGAAGTTTATCGCGTCAAGAAA	1381
Qy	483	PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLys	502
Db	1382	TTCCAGTTTACGTCGAGACATGAGGATGAGGACAGCGACCTCAAGAGGGGGGGAAG	1441
Qy	503	LysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSer	522
Db	1442	AAGCGCTTGGGACATTTGACGAGGACACCCCTCTCTGCTGCTGCTCACCCTCCCAACAGC	1501
Qy	523	SerTrpAenCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCys	542
Db	1502	TCCTGGAACTCGCAGCGGGAGGTCTGCACAGCCCTGCTCATCGAGGTCCAGTCCACTGC	1561
Qy	543	GlnLeuValArgLeuPheAlaArgGlyIleGluGluAenProLysProAspSerHisSer	562
Db	1562	CAGCTGGTTCGACTCTTTGACGAGGAATTTGAAGAGAATCCGAGCCAGACTCACAGC	1621
RESULT 7			
BC067255		1772 bp	linear
LOCUS		1772 bp	linear
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REMARK			
COMMENT			

gene  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6185601"  
/tissue\_type="Peripheral Nervous System, dorsal root ganglion"  
/clone\_lib="Lupski\_dorsal\_root\_ganglion"  
/lab\_host="DH108"  
/note="Vector: PCMV-SPORT6"  
1..1772  
/gene="CERK"  
/note="synonyms: KIAA1646, hCERK, dA59H18.3, DKFP434E0211, FLJ21430, FLJ23239, dA59H18.2, LK4"  
/db\_xref="LocusID:64781"  
579..1598  
/gene="CERK"  
/codon\_start=1  
/product="CERK protein"  
/protein\_id="AAH67255.1"  
/db\_xref="GI:45595583"  
/db\_xref="LocusID:64781"  
/translation="MFSEVILHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGSTD  
CVYSTVGTDAEYTSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGYFGYDIKDS  
EKKRWLGALDYDFSLGKTFVLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGFCVCRQSK  
QOLEEOKALYGLBAEDVREMOVVCGKFLAINATNMSACRRSPRGLSPAHHLDGK  
SSDLILRKSRFNLFLRIHTNQDDQDFTEVEVVKFQFTSKHWEDESDLKE  
GGKRFPHICSHSPSCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIENPK  
PDSHS"

CDS  
Alignment Scores:  
Pred. No.: 1,238-183 Length: 1772  
Score: 2792.50 Matches: 530  
Percent Similarity: 95.15% Conservative: 0  
Best Local Similarity: 95.15% Mismatches: 1  
Query Match: 92.31% Indels: 27  
DB: 9 Gaps: 2

ORIGIN

US-10-631-958-11 (1-562) x BC067255 (1-1772)  
QY 7 ProAlaProLeuGlyValArgAlaProAlaTrpArgThrSerProAlaAlaGluMet 26  
DB 3 CGCGCGCCCTCGCGCTCGCGCGCCCGCCCGCTGGCGGAGAGCCCGCGCGGAGATG 62  
QY 27 GlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValValGlnArgCys 46  
DB 63 GCGGCGACGCGGCGCGGAGCGCTGCAATCCGTCTGCTGGGTGAAGCAGCAGCGCTGC 122  
QY 47 AlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyProGly 66  
DB 123 GCGGTGAGCTGGAGCCCGCGCGGCTCTGCTGCGCTGGTGGCGAGCCCGCGCCGGA 182  
QY 67 AlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaValGlu 86  
DB 183 GCGCGGCGCCCGCGCGGATGCTGCTGCTGCTGTATCTGAGATCATCGCGCTTGAG 242  
QY 87 GluThrAspValHisGlyHisGlnGlySerGlyLysTrpGlnLysMetGluLysPro 106  
DB 243 GAAACAGACGTTTCACGGGAAACATCAAGCAGTGGGAAATGGCAAAATGAAAGCCT 302  
QY 107 TyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTyrTrpAlaGln 126  
DB 303 TACGCTTTTACAGTTTACATGTGTAAGAGAGACACAGCGCACCGCTGGAAAGTGGCGCAG 362  
QY 127 ValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArgGlu 146  
DB 363 GTGACTTCTGTGTCCAGAGAGAGAGCTGTGTCTGTGTGCTGCAGACCTTCGCGGAG 422  
QY 147 MetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPheGly 166  
DB 423 ATGCTGGAGAACTGAGCTGCCAGACCAAGCATTTTACTGGTATTATCAAC----- 473  
QY 167 GlyLys-GlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeuAl 186

DB 474 ---AAAGGACAAAGCGGATATATGAAGAAAGTGGCACCATCTGTTCACTTACG 530  
QY 186 aSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeuTy 206  
DB 531 CTCCATCACCACCTGACATCGC----- 554  
QY 206 rGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPheSe 226  
DB 555 -----ATCGTCTGTGTCGGCGAGATGTTATGTTTCAAG 586  
QY 226 rGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHi 246  
DB 587 CGAGGTGCTGACCGGTCTGATTGGAGGACGACGAGGAGCGCGGGGTTCGACCAACCA 646  
QY 246 sProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySerTh 266  
DB 647 CCCCCGGGCTGTGCTGGTCCCAGTAGCTCCGGATTGGAATCATTCCTCCGACGGGTCAAC 706  
QY 266 rAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisTrl 286  
DB 707 GGACTGCGTGTGTACTCCACCGTGGGCACGACGACGAGAAACCTCGCGCTGCATAT 766  
QY 286 eValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeuLe 306  
DB 767 CGTTGTTGGGACCTCGCTGGCATGGATGTCTCTCAGTCCACCAACACGACACACTCT 826  
QY 306 uArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGl 326  
DB 827 TCGTACTCCGTGTCCTGCTGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGA 886  
QY 326 uLysIleArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSe 346  
DB 887 GAAGAAACCGGTGGTGTGGCTGTTCAGATACGACATTTTAAAGAGCTTCTCTCTC 946  
QY 346 rHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerPr 366  
DB 947 CCACACTGCTATGAGGACAGTGTCTTCTCTCTCCGTGCACACACACGCTGGGATCTCC 1006  
QY 366 oArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLe 386  
DB 1007 AAGGATAGGAAGCCCTGCGCGGAGGATGCTTTGTTGACGGAACAAAGCAGCAGCT 1066  
QY 386 uGluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluTr 406  
DB 1067 GGAGGAGGAGCAGAGAAAGCACTGTATGGTGTGGAGCTCGGAGGACGTGGAGGAGTG 1126  
QY 406 pGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysAr 426  
DB 1127 CAAAGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCTGTGCTTGTG 1186  
QY 426 gArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIl 446  
DB 1187 CCGGAGCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCTGACCTCAT 1246  
QY 446 eLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsnGl 466  
DB 1247 CCTCATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAACCAACCA 1306  
QY 466 nGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPheTh 486  
DB 1307 CGAGGACAGTTTGACTTCACTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGATTTAC 1366  
QY 486 rSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPheGl 506  
DB 1367 GTCGAAGCACATGAGGATGAGGACGACCTCAAGGAGGGGGGGAAGAGCGCTTTGG 1426  
QY 506 yHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsnCy 526  
DB 1427 GCACATTTGACAGCAGCCCTCTCTGCTGTGTCACGCTCTCCAAACAGCTCTCTGGAAC 1486  
QY 526 sAspGlyValValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuValAr 546  
DB 1487 CGAGGGGAGGTCTCTGCACAGCCCTGCCATCGAGGTGAGATGCCACTGCCAGCTGGTTCG 1546

DEPT. OF

```

ACCESSION   BD183468
VERSION     BD183468.1  GI:31875668
KEYWORDS    JP 2002345492-A/181.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 4171)
AUTHORS     Ohara, O., Nagase, T. and Nakajima, D.
TITLE       Novel genes and proteins encoded by the genes
JOURNAL     Patent: JP 2002345492-A 181 03-DEC-2002;
            KAZUSA DNA RESEARCH INSTITUTE
COMMENT     OS Homo sapiens (human)
            PN JP 2002345492-A/181
            PD 03-DEC-2002
            PF 26-FEB-2002 JP 2002049009
            PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA
            PC C12N15/09, C07K14/47//A61K31/711, A61K38/00, A61P25/00,
            A61P25/14,
            PC A61P25/19, A61P35/00, C12N15/00, A61K37/02
            CC Novel genes and proteins encoded by the genes FH Key
            Location/Qualifiers
            FT CDS
               (2)..(1444).
FEATURES             Location/Qualifiers
     source           1..4171
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.:          4,848-169      Length:      4171
Score:              2588.00        Matches:    481
Percent Similarity: 100.00%        Conservative: 0
Best Local Similarity: 100.00%      Mismatches:  0
Query Match:        85.55%          Indels:      0
DB:                 6              Gaps:        0

US-10-631-958-11 (1-562) x BD183468 (1-4171)
QY 82 ileileAlaValGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGln 101
DB 2 ATCATCGCGTGTAGGAAACAGACGCTTACGGGAAACATCAAGGAGTGGAAATGGCAG 61
QY 102 LysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgHisArg 121
DB 62 AAAATGGAAAAGCCCTTACGCTTTTACAGTTTCACTGTGTAAGAGAGACGACGCGCCGC 121
QY 122 TrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeu 141
DB 122 TGGAAAGTGGCGCGAGGTGACTTTCTGTTGTCCAGAGGAGCAGCTGTGTCACTTGTGGCTG 181
QY 142 GlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPhe 161
DB 182 CAGACCCCTGCGGGAGATGCTGGAAGCTGAGCTCCAGACCAACATTTACTGGTATTT 241
QY 162 ileAsnProPheGlyGlyGlnGlyLysArgGlleTyrGluArgLysValAlaPro 181
DB 242 ATCAACCCGTTTGGAGGAAAGCAAGCAAGCGGATATATGAAGAAAGTGGCACCA 301
QY 182 LeuPheThrLeuAlaSerIleThrAspIleIleValThrGluHisAlaAsnGlnAla 201
DB 302 CTGTTTCACTTACCTCCATCACCCTGACATCATCTGTTACTTGAACATGCTAATCAGGCC 361
QY 202 LysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGly 221
DB 362 AAGGAGACTCTGTATGAGATTAACTAGACAATACGCGCATCTGTGTGTGCGCGGA 421
QY 222 AspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGly 241
DB 422 GATGTATGTTTACGCGAGGTGCTGCACGCTCTGATTGGAGACCGCAGAGGCGCCGGG 481
QY 242 ValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIle 261

```

```

Db 482 GTCCACCAAGAACACCCCGGGCTGTGTGTGTCCTCCAGTAGCCTCCGATTGGAAATCATTT 541
QY 262 ProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThr 281
Db 542 CCCGACGGGTCAACGGACTGCTGTGTGTACTCCACCGTGGGCACCGACGCGCAGAAACC 601
QY 282 SerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHis 301
Db 602 TCGCGCTGTCATATCGTTGTTGGGACTCGCTGCCCATCGATGTGTCTCTAGTCCACCAC 661
QY 302 AsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIle 321
Db 662 AACAGCACATCTCTCGCTACTCTCGTGTCCCTGTGGCTACGGCTTCTACGGGACATC 721
QY 322 IleLysAspSerGluLysLysLeuArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeu 341
Db 722 ATCAAGGACACGTAGAGAAACGGTGGTGGTCTTTCGCAGATACGACTTTTCAGGTTTA 781
QY 342 LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis 361
Db 782 AAGACCTTCTCTCCACCACTGCTATGAAGGGACAGTGTCTCTCTCTCCGACAAACAC 841
QY 362 ThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGln 381
Db 842 ACGTGGGATCTCCAAGGGATAGGAAGCCCTGCCGGCAGGATGCTTTTGTTCAGGCAA 901
QY 382 SerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGlu 401
Db 902 AGCAAGCAGCAGCTGGAGGAGGACAGAAAGACACTGTATGTTTGAAGCTGCGGAG 961
QY 402 AspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMet 421
Db 962 GACGTGGAGGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCCCAAAACATG 1021
QY 422 SerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGly 441
Db 1022 TCTGTGCTTGTGCGCGAGGCCCCAGGGGCTCTCCCCGGCTGCCCACTTGGAGACGGG 1081
QY 442 SerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIle 461
Db 1082 TCTTCTGACCTCATCTCATCCGAAATGCTCCAGGTTCAATTTTCTCAGATTTCTCATC 1141
QY 462 ArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLys 481
Db 1142 AGGCACACCAACACGACGACGACGTTTGTGACTTCACTTTTGTGAAGTTTATCGCGTCAAG 1201
QY 482 LysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyGly 501
Db 1202 AAATTCCAGTTTACGTGAAGCACATGGAGATGAGGACGACGACCTCAAGAGGGGGGG 1261
QY 502 LysLysArgPheGlyHisIleCysSerSerHisProSerCysCysCysThrValSerAsn 521
Db 1262 AAGAAAGCGCTTGGGCACATTTGGACGACGACACCCCTCTCTGCTGCTGACGCTCCAAAC 1321
QY 522 SerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHis 541
Db 1322 AGCTCCTGGAACCTCGACGGGAGGTCTGTCACAGCCCTGCTGCTGACGAGTCCAGTCCAC 1381
QY 542 CysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHis 561
Db 1382 TGGCAGCTGTTGTCAGCTCTTTTGCACGAGGAATGAAGAGAATCCGAAGCCAGACTCACAC 1441
QY 562 Ser 562
Db 1442 AGC 1444

```

```

RESULT 10
AB051433 LOCUS
DEFINITION Homo sapiens mRNA for KIAA1646 protein, partial cds.
ACCESSION AB051433
VERSION AB051433.1 GI:13359166

```



Db 1322 ACCTCTGGAAGTGGAGCGGGAGGTCTGTCACACCCCTGCCATCGAGTCAAGTCCAC 1381  
Qy 542 CysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAAsnProGlyProAspSerHis 561  
Db 1382 TGCCAGCTGGTTCGACTCTTTTCACGAGGAATTGAAGAGATCCGAGCCAGACTCACAC 1441  
Qy 562 Ser 562  
Db 1442 AGC 1444  
RESULT 11  
CQ730476  
LOCUS CQ730476 1459 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 16410 from Patent WO02068579.  
ACCESSION CQ730476  
VERSION CQ730476.1 GI:42304409  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.  
AUTHORS  
TITLE Kits, such as nucleic acid arrays, comprising a majority of  
humanexons or transcripts, for detecting expression and other uses  
thereof  
JOURNAL Patent: WO 02068579-A 16410 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES  
Location/Qualifiers  
source 1..1459  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,176-164 Length: 1459  
Score: 2510.00 Matches: 478  
Percent Similarity: 87.89% Conservative: 1  
Best Local Similarity: 87.71% Mismatches: 6  
Query Match: 82.98% Indels: 61  
DB: 6 Gaps: 3  
US-10-631-958-11 (1-562) x CQ730476 (1-1459)  
Qy 18 TrpArgThrSerProAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSer 37  
Db 3 TGGCGGACGAGCCCGCGCGGAGATGGGGCGACGGGGCGCGGAGCCGCTGCANACC 62  
Qy 38 ValLeuTrpValLysGlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeu 57  
Db 63 GTGCTGTGGGTGAAGCAGCAGCGCTGCCCAA----- 94  
Qy 58 ArgTrpTrpArgSerProGlyProGlyAlaGlyAlaProGlyAlaAspAlaCysSerVal 77  
Db 95 -----GGACGTGGTGT-----TTTACAGATGCCTGTCTGTG 127  
Qy 78 ProValSerGluIleAlaValGluThrAspValHisGlyLysHisGlnGlySer 97  
Db 128 CTTGATCTGAGATCATCGCGGTTGAGGAAACAGACGTTTACGGTTTACAGTTCACTGTGTAAGAGAGCA 187  
Qy 98 GlyLysTrpGlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArgAla 117  
Db 188 GGAAATGGCAGAAATGMAAGCCTTACGCTTTTACAGTTCACTGTGTAAGAGAGCA 247  
Qy 118 ArgArgHisArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGlnLeuCys 137  
Db 248 CGACGGCACCCTGGAAAGTGGCGCGAGGTGACTTTCTGTGTCTCCAGAGGAGCAGCTGTGT 307  
Qy 138 HisLeuTrpLeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHis 157  
Db 308 CACTGTGTGGCTGCAGACCCCTGCGGAGATGTGGAGAGCTG----- 349

Qy 158 LeuLeuValPheIleAAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArg 177  
Db 349 ----- 349  
Qy 178 LysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHis 197  
Db 350 -----ATTACTGAACAT 361  
Qy 198 AlaAsnGlnAlaLysGluThrLeuTyrGluIleAAsnIleAspLysTyrAspGlyIleVal 217  
Db 362 GCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCGTC 421  
Qy 218 CysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLysGlyArgThrGln 237  
Db 422 TGTGTGCGCGGAGATGGTATGTTTTCAGCGAGGTGTCACCGGTCTGATTGGGAGGACGAG 481  
Qy 238 ArgSerAlaGlyValAspGlnAAsnHisProArgAlaValLeuValProSerSerLeuArg 257  
Db 482 AGGAGCGCGGGGTGCGACCAACACCCCGGGCTGTGCTGCTCCCGTAGCTCCGG 541  
Qy 258 IleGlyIleIleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSer 277  
Db 542 ATTGGAATCATTTCCCGCAGGGTCAACGGACTGCGTGTGTACTCCACCGTGGCACACGAGC 601  
Qy 278 AspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSer 297  
Db 602 GACGCAGAAACCTCGCGGCTGCATATCTGTTGGGAGCTCGCTGGCCATGGATGTGTC 661  
Qy 298 SerValHisHisAAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPhe 317  
Db 662 TCAGTTCACCAACAACAGCAGACTCTCTTCGCTACTCCCGTGTCTGCTGGGCTACGGCTTC 721  
Qy 318 TyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLysAlaAlaArgTyrAsp 337  
Db 722 TACGGGACATCATCAAGGACAGTGAGAAGAAACGGTGTGTTGGTCTTGCACAGATACGAC 781  
Qy 338 PheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeu 357  
Db 782 TTTTTCAGGTTTAAAGACCTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCCTC 841  
Qy 358 ProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPhe 377  
Db 842 CCTGCACAACAACACGGTGGGATCTCCAAGGATAGGAAGCCCTGCCGGGAGGATGCTTT 901  
Qy 378 ValCysArgGlnSerLysGlnGlnLeuGluGlnLysLysAlaLeuTyrGlyLeu 397  
Db 902 GTTTGCAGGCAAAACAGCAGCAGCTGGAGGAGGAGCAGAGAAAGCACTGTATGTTTG 961  
Qy 398 GluAlaAlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAAsn 417  
Db 962 GAAGCTCGGAGGACGTGGAGGAGTGGCAAGTCTGTGTGGGAAAGTTTCTGGCCATCAAT 1021  
Qy 418 AlaThrHisMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHis 437  
Db 1022 GCCAAACATGTCTCTGTGCTTGTGCGGAGCCCGAGGGGCTCTCCCGGCTGCCAC 1081  
Qy 438 LeuGlyAspGlySerSerAspLeuIleLeuLysCysSerArgPheAAsnPheLeu 457  
Db 1082 TTGGGAGACGGGTCTTCTGACCTCATCTCATCGGAAATGCTCCAGGTTCAATTTCTG 1141  
Qy 458 ArgPheLeuIleArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluVal 477  
Db 1142 AGATTTCTCATCGGCACACCAACAGCAGGACAGTTTGTACTTCTCCTTTTGTGAAGTT 1201  
Qy 478 TyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeu 497  
Db 1202 TATCGCGTCAAGAAATTCAGTTTACGTGCAAGACATGAGGAGATGAGAGCAGCGACCTC 1261  
Qy 498 LysGluGlyLysLysArgPheGlyHisIleCysSerSerHisProSerCysCys 517  
Db 1262 AAGGAGGGGGGGAAGAGCGCTTTTGGGCACATTTGACAGCACCCCTCTCTGCTGCTGC 1321



```

Qy 518 ThrValSerAenSerSerTrpAenCysAspGlyGluValLeuHisSerProAlaIleGlu 537
Db 1322 ACCGTCCTCAACAGCTCTCTGAACTGCGACGGGAGGTCTCTGCACAGCCCTGCCATCGAG 1381

Qy 538 ValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAenProIlys 557
Db 1382 GTACAGATCCACTCCCACTGCTGTTTCGACTCTTTCCACGAGGAATGAAGAAATCCGAAG 1441

Qy 558 ProAspSerHisSer 562
Db 1442 CCAGACTCACACAC 1456

RESULT 12
AB079067 2830 bp mRNA linear ROD 27-JUN-2002
LOCUS AB079067 Mus musculus cerk mRNA for ceramide kinases, complete cds.
DEFINITION AB079067
ACCESSION AB079067
VERSION AB079067.1 GI:21624341
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Sugiura, M., Kono, K., Liu, H., Shimizugawa, T., Minekura, H.,
Spiegel, S., and Kohama, T.
Ceramide Kinase, a Novel Lipid Kinase. MOLECULAR CLONING AND
FUNCTIONAL CHARACTERIZATION
J. Biol. Chem. 277 (26), 23294-23300 (2002)
11956206
2 (bases 1 to 2830)
Sugiura, M., Kono, K., Shimizugawa, T., Minekura, H., Spiegel, S. and
Kohama, T.
Direct Submission
Submitted (29-JAN-2002) Masako Sugiura, Sankyo Co., Ltd.,
Pharmacology and Molecular Biology Research Laboratories; 2-58
Hiromachi 1-chome, Shinagawa-ku, Tokyo 140-8710, Japan
(E-mail: msugiura@shina.sankyo.co.jp, Tel:81-3-3492-3131,
Fax:81-3-5436-8565)
FEATURES
Location/Qualifiers
1..2830
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..2830
/gene="cerk"
1..1596
/gene="cerk"
/codon_start=1
/product="ceramide kinases"
/protein_id="BAC01155.1"
/db_xref="GI:21624342"
/translation="MGAMGAEPHLSVLWVRRRCVLSLEPARALLRWRSPPGPSPA
PGADARSLVSEIIAIVEKDDCEKSHSGRWKHNPFATVHRVKVRHHRWKWARV
TFWSADQLCHLQTLRLGLESLSRSPKHLVFPNGGKGQGRKYIEKTVAPLFTL
ASITETIITTHANOAKETLYEINTDSYDGVICVCGDGMFSEVLHGVIRTOQSAGID
PNHPRVLVSTLRIGIIPAGSTDVCYSTVGTNDAAETSAHHIIIGDPLADVSSVHY
QHTLLRSVLLGCGFGVDLLIKDSEKRWMLVRYDFSGLKTFLSHQYVYEGTLSFLPA
QHTVGSPRNDKPCAGCFVCKQSKQLEEEKALYGLIENAEVEEQVTCGRFLIN
ATNMSCAPRSPGSLSPFAHLGDSGLIIRKCSRFNLFRLIRHTNQDQDFTFV
EYVRVKFHFSTSKVEDEDNDSDKEQKFGKIKDRPSCCTCSASRSSWNGDGEVMS
PAIEVRVHCQLVRLFARGIEES"
ORIGIN
Alignment Scores:
Pred. No.: 8.72e-159 Length: 2830
Score: 2436.50 Matches: 454
Percent Similarity: 90.19% Conservative: 33
Best Local Similarity: 84.07% Mismatches: 48
Query Match: 80.55% Indels: 5
Db: 10 Gaps: 1
US-10-631-958-11 (1-562) x AB079067 (1-2830)

```

```

Qy 26 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValIysGlnArg 45
Db 1 ATGGGGCAATGGGGCGGGAGCGCTGCACCTCCGTGCTGTGGGTGAAACGGCGAGCGC 60

Qy 46 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 65
Db 61 TGTGCGCTCAGCTGGAGCCCGGGGCTCTGTACGCTGGTGGCGGAGCCCGGAGGCC 120

Qy 66 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleIleAlaVal 85
Db 121 GGGCCCTCCGCGGGGTGCGATGCCCTCGGTACTGGTTCGGAGATCATCGCTGTT 180

Qy 86 GluGluThrAspValHisGlyHisGlnGlySerGlyLysTrpGlnLysMetGluLys 105
Db 181 GAGGAAAAGACGACTCGGAGAAACACGCCCTCCAGTGGCCGATGATATAAATGAAAAAT 240

Qy 106 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAla 125
Db 241 CCGTTCGATTTCAGTCCACCGTGTGAAGCGAGTACGACATCACCGCTGGAGTGGGCA 300

Qy 126 GlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 145
Db 301 CGGCTGACCTTCTGGAGCGCGAGCAGCTGTGTACCTGTGGCTGCAGACCTCCGT 360

Qy 146 GluMetLeuGluLysLeuThrSerArgProLysHisLeuValPheIleAsnProPhe 165
Db 361 GGGCTGCTGGAGAGCCCTGACTTCAAGACCGAGCATTGCTGGTATTTCATCAACCTTTC 420

Qy 166 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 185
Db 421 GGAGGAAAGGTGAGGCAAGCGCATCTATGAAAAAAGACGTGGCGCTCTGTTTACCTTG 480

Qy 186 AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 205
Db 481 GCTTCCATCACTACGGAGATCATCATTTACAGAGCATGCCAACCAAGCAAGAGACATTTA 540

Qy 206 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 225
Db 541 TACGAGATCAACACAGACAGCTATGATGSCATCGTGTGCGTGGTGGGAGCGCATGTTT 600

Qy 226 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 245
Db 601 AGCGAGGTGCTGCATGGGTGCTATGGGAGGACGAGCAGCGCTGGTATCGACCCCAAT 660

Qy 246 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 265
Db 661 CACCCCGAGCGGTGCTGGTGGCCCGAGTACCTTCAGGATCGGCATCATACCCGCGAGGTCC 720

Qy 266 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 285
Db 721 ACAGATTGTGTGTTACTCAACAGTAGGCACAAACCCAGCAGACATCGGCTTTGCAT 780

Qy 286 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 305
Db 781 ATCATTTATGGGACTCACTGGCAATAGAGTGTCTCTGTGCATCTACCAATACCGCTG 840

Qy 306 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer 325
Db 841 CTGCGGTACTCGGTTTCTCTGCTGGGTACGGTTTCTACGGGGAGCTTAATCAAGACAGC 900

Qy 326 GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 345
Db 901 GAAAAGAAACGGTGGATGGGCTCTGTCGGTATGATTTCTCAGGGTTGAAGACCTTTCTC 960

Qy 346 SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 365
Db 961 TCTCATGACTATGAAGGAGACACTGTCTCTCCAGCAGACACACCGTGGGATCT 1020

Qy 366 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 385
Db 1021 CCACGGCAACAATAAACCTCTGCCGGCTGGGTGCTTGTGTGCGAGCAGACAGACACAG 1080

```



```
Qy 386 LeuGluGluGluGluGlyLysAlaLeuTyTGlyLeuGluAlaAlaGluAspValGluGlu 405
Db 1081 CTGGAAGAAGAAAGAAAGAAACCTGTATGGCCCTGGGAGAACCGCGAGGAAGTGAAGAG 1140
Qy 406 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 425
Db 1141 TGGCAAGTGACATGTGGGAGTTCTTGCCCATCATGCCACCAACATGTCCTGTCTGTGT 1200
Qy 426 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 445
Db 1201 CCTCGAGCCCTGGGGGCGCTGTCCCATTTGGCCCATTTGGCCCATCTGGGAGATGGGTCTTGACCTC 1260
Qy 446 IleLeuIleAArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 465
Db 1261 ATCTCTATCCGAAGTGTCTCAGGTCAACTTCTTGAGATTCTCATCCGCACACGAAC 1320
Qy 466 GlnGlnAspGlnPheAspPheThrPheValGluValTyArgValLysLysPheGlnPhe 485
Db 1321 CAGGAGGACCATTCGACTTCACTTTCGTTGAGTTTATCGAGTCAAGAAATTCACCTTC 1380
Qy 486 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe 505
Db 1381 ACGTCAAGACAGTGGAGATGAGGACAAATGACTCGAAGGAACAAGAGAGAGAGGTTT 1440
Qy 506 GlyHisIleCysSerSerHisProSerCysCysCysThrValSerAsnSerSerTrpAsn 525
Db 1441 GGAAGATCTTGAAGACAGACCCCTCTTGCACCTTCCTCAGGCTCCAGAAAGCTCTGGAAC 1500
Qy 526 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 545
Db 1501 TCGATGCGGAAGTATGCACAGCCCGGCCATTGAGGTTCAGGTTCATCGCAGCTGGTG 1560
Qy 546 ArgLeuPheAlaArgGlyIleGluGlu-----AsnProLysProAspSer 560
Db 1561 GCCTCTTTTGTCTCGGGGAATCGAGGAGAGTCAATAGCAAGAACCCCAAGCCAGGAGC 1620

RESULT 13
AK129416
LOCUS Mus musculus mRNA for mKIAA1646 protein. linear ROD 21-NOV-2003
DEFINITION AK129416
ACCESSION AK129416.1 GI:37360495
VERSION FLI_CDNA.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Koseki,H., Hiraoka,S.,
Saga,Y., Nagase,T., Ohara,O. and Koga,H.
TITLE Prediction of the coding sequences of mouse homologues of KIAA
gene: III. the complete nucleotide sequences of 500 mouse
KIAA-homologous cDNAs identified by screening of terminal sequences
of cDNA clones randomly sampled from size-fractionated libraries
JOURNAL DNA Res. 10 (4), 167-180 (2003)
MEDLINE 22977043
PUBMED 14621295
REFERENCE 2 (bases 1 to 3661)
AUTHORS Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.
TITLE Direct Submission
JOURNAL Submitted (23-JUN-2003) Hisashi Koga, Kazusa DNA Research
Institute, Laboratory for Genome Informatics, 2-6-7
Kazusa-kanatari, Kisarazu, Chiba 292-0818, Japan
(E-mail:mouse@kazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)
COMMENT The CREATE program supported by Japan science and technology
corporation; cDNA full insert sequencing; Kazusa DNA Research
Institute; cDNA library construction, clone selection and 5'- &
3'-end one pass sequencing.
LOCATION/Qualifiers
FEATURES
1..3661
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
```

```
gene
CDS
/clone="mph01246"
/tissue_type="embryonic tail"
/notes="vector:modified pBC SK+"
1..3661
/gene="mKIAA1646"
<1..1231
/gene="mKIAA1646"
/notes="CDS is predicted by in silico analysis. Start codon
is not identified.
Start codon is not identified."
/codon_start=2
/evidence=not experimental
/product="mKIAA1646 protein"
/protein_id="BAC398226.1"
/db_xref="GI:37360496"
/translation="FCFAASRPKHLVFINPFGGQGGKRIYEKTVAPLFTLASITTE
III TEHANOAKETLYEINTDSYDGI VCVGGDMFSEVLHGIVIGTQOSAGIDPNHPRA
VLVPTLRIGIIPAGSDTCVYSTVGTNDATSAHLII IGDLSLAIDSSVHYHNTLLR
YSVLLGYGPGYGLIDKSEKRWMLVRFDSGLKTFLSHQYVYEGTSLFSLAQNHTVGS
PRDNCRAGCFVCRSQQLLEBEKKALYGLNAEEVEEQVTCGKFLAINATMSC
ACPRSPGLSPFAHLGDGSDLIIRKCSRPNFLRFLIRHTNQDQDFDTFVYVRVK
KFHPTSKHVEDENDNDSKEQKQKFGKICKORPSCSTCSASRSSMNCDEGVMHSPAIEVR
VHCQLVRLFARGIEES"
```

ORIGIN

Alignment Scores:

Pred. No.:	1,92e-122	Length:	3661
Score:	1912.50	Matches:	356
Percent Similarity:	92.49%	Conservative:	26
Best Local Similarity:	86.20%	Mismatches:	26
Query Match:	63.22%	Indels:	5
DB:	10	Gaps:	1

US-10-631-958-11 (1-562) x AK129416 (1-3661)

Qy 153 SerArgProLysHisLeuLeuValPheIleAsnProPheGlyGlyGlyGlnGlyLys 172

Db 17 TCAAGACGGAAGCACTTGTCTGGTATTATCAACCCCTTCGAGGAGAAAGGTCAAGGCAAG 76

Qy 173 ArgIleTyArgGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIle 192

Db 77 CGCATCTATGAAACAAACAGTGGCGCCTCTGTTTACCTTGGCTTCATCCTACGAGATC 136

Qy 193 IleValThrGluHisAlaAsnGlnAlaLysGluThrLeuTyArgGluIleAsnIleAspLys 212

Db 137 ATCATTCAGACAGCATGCCAACCAAGCCAAAGAGAGACTTTATACGAGATCAACACACAGCAGC 196

Qy 213 TyrAspGlyIleValCysValGlyGlyAspGlyMetPheSerGluValIleHisGlyLeu 232

Db 197 TATGATGGCATCGTGTGGTAGGTGGGACGGCATGTTTCAGCGAGGTCTCGATGGGGTG 256

Qy 233 IleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuVal 252

Db 257 ATTGGGAGGAGCAGCAGAGCGCTGGTATCGACCCCAATCACCCCGAGCGGTCTGTGTG 316

Qy 253 ProSerSerLeuArgIleGlyIleProAlaGlySerThrAspCysValCysTySer 272

Db 317 CCCAGTACCCCTCAGGATCGGCATCATATCCCGAGGGTCCACAGATTGTGTGTGTACTCA 376

Qy 273 ThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeu 292

Db 377 ACATGGGCAACAAACGACGACATCGGCTTTGCACATCATTTATTTGGGAGCTCGGTG 436

Qy 293 AlaMetAspValSerSerValHisHisAsnSerThrLeuLeuArgTySerValSerLeu 312

Db 437 GCATAGACGTGTCTCTGTGCATACCATTAACAGGTGCTGCGGTACTCGGTTCCTCTG 496

Qy 313 LeuGlyTyArgPheTyArgLysAspIleIleLysAspSerGluLysLysArgTyLeuGly 332

Db 497 CTGGGCTACGGTTCTACGGGGACTTAATCAAGACAGCGAAGAAAGAAACCGGTGATGGGC 556

Qy 333 LeuAlaArgTyArgPheSerGlyLeuLysThrPheLeuSerHisHisCysTyTyGluGly 352

```
Db 557 CTGCTCCGGTATGATTTCTCAGGTTGAAGACCTTTCTCTCTCATCTAGTACTATGAAGG 616
Qy 353 ThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysProCys 372
Db 617 ACCTGTCTTCTCCACAGCAGACACGGTGGATCTCCACGGGACAATAAACCTTGC 676
Qy 373 ArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLeuGluGluGlnLysLys 392
Db 677 CGGCTGGTCTCGTGTGAGGAGAGCAAGCAACAGCTGGAGAGAGAGAGAGAA 736
Qy 393 AlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGluTrpGlnValValCysGlyLys 412
Db 737 GCCCTGTATGCTGGAGAGCCCGAGAGAGTGAAGAGTGGCAAGTGCATGTGGGAAG 796
Qy 413 PheLeuAlaIleAenAlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeu 432
Db 797 TTCTTGGCATTCAATGCCACCACATGCTCTGTCTGTCTCGAGCCCTGGGGGCTG 856
Qy 433 SerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIleLeuIleArgLysCysSer 452
Db 857 TCCCAATTGGCCATCTGGGAGATGGTCTTCTGCACCTCATCTTATCGGAAGTGTCTCC 916
Qy 453 ArgPheAenPheLeuArgPheLeuIleArgHisThrAenGlnGlnAspGlnPheAspPhe 472
Db 917 AGGTTCAACTTCTCAGATTTCTCATCCGGCACACGAACCAAGAGAGACCACTTCGACTTC 976
Qy 473 ThrPheValGluValTyrArgValLysLysPheGlnPheThrSerLysHisMetGluAsp 492
Db 977 ACTTTCGTGTAAGTTATCGAGTCAAGAAATTCACCTTCAGCTCGAAGCAGTGGAGAC 1036
Qy 493 GluAspSerAspLeuLysGluGlyGlyLysArgPheGlyHisIleCysSerSerHis 512
Db 1037 GAGGACAATGACTCGAAGGACAGAGAGAGCAAGAGTTTGGGAAGATCTCGAAGACAGA 1096
Qy 513 ProSerCysCysThrValSerAsnSerSerTrpAenCysAspGlyGluValLeuHis 532
Db 1097 CCCTCTTGCATCTGCTCAGCCTCCAGAGCTCTCGAAGCTCGCAGCGGGAGTGCAC 1156
Qy 533 SerProAlaIleGluValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIle 552
Db 1157 AGCCCGGCATTGAGTCAAGGTTCATCGCAGCTGTCAGCTGTGCGCCCTTTGCTCGGGGATC 1216
Qy 553 GluGlu-----AsnProLysProAspSer 560
Db 1217 GAGGAAGAGTCATAAGCAAGAACCACCAAGGCCAGGAGC 1255

RESULT 14
CR386590 1450 bp mRNA linear VRT 05-APR-2004
LOCUS Gallus gallus finished cdna, clone ChEST291d19.
DEFINITION CR386590
ACCESSION CR386590.1 GI:46239349
KEYWORDS Gallus gallus (chicken)
SOURCE Gallus gallus
ORGANISM Gallus gallus
REFERENCE 1 (bases 1 to 1450)
AUTHORS Boardman,P.E., Bonfield,J.K., Brown,W.R.A., Carder,C., Chalk,S.E.,
Croning,M.D.R., Davies,R.M., Francis,M.D., Grafham,D.V.,
Hubbard,S.J., Humphray,S.J., Hunt,P.J., Maddison,M., McLaren,S.R.,
Niblett,D., Overton,I.M., Rogers,J., Scott,C.E., Taylor,R.G.,
Tickle,C. and Wilson,S.A.
DIRECT SUBMISSION
SUBMITTED (05-APR-2004) Sanger Institute, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: chickens@bms.umist.ac.uk
BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cdna
sequencing project.
This sequence is from the
BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cdna collection,
from a library constructed by Elizabeth Bosch. cdna was prepared
from RNA extracted from heads, normalised, and poly A-trimmed.
```

```
FEATURES
source
1..1450
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="ChEST291d19"
/clone_lib="CSEQCHN23"
/dev_stage="stage 22"

ORIGIN
Alignment Scores: 4.9e-115 Length: 1450
Pred. No.: 1798.00 Matches: 322
Score: 83.30% Conservative: 72
Percent Similarity: 68.08% Mismatches: 73
Best Local Similarity: 59.44% Indels: 7
Query Match: 5 Gaps: 3
Db: 3

US-10-631-958-11 (1-562) x CR386590 (1-1450)
Qy 93 LysHisGlnGlySerGlyLysTrpGlnLysMetGluLysProTyrAlaPheThrValHis 112
Db 1 AAACAGCGTAGTATTGGCAAAATGGCAAAAGATGCAAAAGCCCATGCTTTACAGTTTAT 60
Qy 113 CysValLysArgAlaArgArgHisArgTrpLysTrpAlaGlnValThrPheTrpCysPro 132
Db 61 TACGTGAAAAAAGCCCAAAATCACCGCTGGCGGTGCAGAGATGTGAGCTTTGTGTGCT 120
Qy 133 GluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArgGluMetLeuGluLysLeuThr 152
Db 121 GATGAGATCTTGTGTACCACTGATACAGGCACT----GAATTACTTGAATTAATAATAA 176
Qy 153 SerArgProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLys 172
Db 177 TGTAAGCAAAAGCAGTTGCTTGTGTATATTAATCATATGAGGAAAAACCAAGGGAAG 236
Qy 173 ArgIleTyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIle 192
Db 237 AGGATTTATGAACAGAAAGTTGCTCCACTCTTCAGCTTGGCTTCTATTCTTACTGATGTT 296
Qy 193 IleValThrGluHisAlaAsnGlnAlaLysGlnThrLeuTyrGluIleAsnIleAspLys 212
Db 297 GTCGTAACAGCATGCTAACCATGCTAAGACAACTATTGGAAGTTAATAATAATAA 356
Qy 213 TyrAspGlyIleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeu 232
Db 357 TATGATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 416
Qy 233 IleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuVal 252
Db 417 ATTGGAAGATGCAAGAGGACTCTGGCATAGACAAATAATCCCAAGACCGGTAGTC 476
Qy 253 ProSerSerLeuArgIleGlyIleProAlaGlySerThrAspCysValCysTyrSer 272
Db 477 CAGTGCATATAGCATGGCATTAATTCCTGCTGGCTCAACAGATTCGGTATGCTATTC 536
Qy 273 ThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeu 292
Db 537 ACTGTGGCATTTCTGATCTGATCAACATCAGCTCTTCATATTTATATAGGTGACTGTG 596
Qy 293 AlaMetAspValSerSerValHisAsnSerThrLeuLeuArgTyrSerValSerLeu 312
Db 597 CCTTTAGATGTCATCTGTGCATCAAAACACACGTTTTTGAAGATGCTGTATCATG 656
Qy 313 LeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGly 332
Db 657 TTGGGCTATGTTTTTATGGAGATGTTTTGAAAGACAGTGAAGAAACGCGTGGATGGT 716
Qy 333 LeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGly 352
```

ECORI-NotI cut cdna was then ligated into the vector. Vector:  
pBluescript II KS(+); Site\_1: EcoRI; Site\_2: NotI Host: Escherichia  
coli DH10B.

```
Db 717 CGATGAGATAGCATATTCAGGCTTCAAGACTTTTCTTCTCATCATATCATGAGA 776
Qy 353 ThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysProCys 372
Db 777 ACAATTTCTTTTCAACCAACGAAACACACACTGGGATCTCCACGAGATAAAGATAGCTGC 836
Qy 373 ArgAlaGlyCysPheValCysArgGlnSerLysGlnLeuGluGluGlnLysLys 392
Db 837 AGAACAGATGTTATTTTGAAGGAAAGTGAAGCAACAACTGGCAGAACAAACGCAAGAAG 896
Qy 393 AlaLeuTyGlyLeuGluAlaAlaGluAspValGluGluTyrGlnValValCysGlyLys 412
Db 897 TGT-----GGATTCAACATGAGAGATGAGAGAGATGAGAGTTATTAAAGGGAAA 950
Qy 413 PheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgGlySerProArgGlyLeu 432
Db 951 TTTCTAGCCATCAATGTAGTAAATATGCTGTGCTCTCCACGAGTCCAAAGTCTT 1010
Qy 433 SerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIleLeuIleArgLysCysSer 452
Db 1011 TCACGGCAGCTCATTTTACAGATGTTTACAGTCACTGACCTGATTTAGTTTCGGAATGCTTCC 1070
Qy 453 ArgPheAsnPheLeuArgPheLeuIleArgHisThrAsnGlnGlnAspGlnPheAspPhe 472
Db 1071 AGATTGATTTTCTGCGTTATCTTGTGACATACAAACAAAGATGACAGTTTGACTTC 1130
Qy 473 ThrPheValGluValTyrArgValLysLysPheGlnPheThrSerLysHisMetGluAsp 492
Db 1131 TCATTTGTAGATGTTTATCGGTTGAAGATTTTCAATTTTACATCAAGCTTTCAGAGAC 1190
Qy 493 GluAspSerAspLeuLysGluGlyCysLysLysArgPheGlyHisIleCysSerSerHis 512
Db 1191 AATGAAAGCAGTGTACACAGATAGAGAAAGAAACATTTTGGCCAGTTCTGCAGAGATCAT 1250
Qy 513 ProSerCysCysThrValSerAsnSerSerTyrPheAspGlyValGluValLeuHis 532
Db 1251 CCAGCTCTCTGCAATATTTGCTAATAGCACTTGAATGTTGATGGAGAAACCTTGAT 1310
Qy 533 SerProAlaIleGluValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIle 552
Db 1311 AGCTCAGCAATTGAAGTGAAGTTCACCTGCCAGTTAATGAACTGTTTGCAGGGGAATC 1370
Qy 553 GluGluAsnProLysProAsp-----SerHisSer 562
Db 1371 GAGGAAACTGTAAAGATGAAGCTGCCTACAGCCACAGT 1409

RESULT 15
BC074350 2494 bp mRNA linear VRT 03-AUG-2004
LOCUS Xenopus laevis MGC84197 protein, mRNA (cDNA clone MGC:84197
DEFINITION IMAGE:6954181), complete cds.
ACCESSION BC074350
VERSION BC074350.1 GI:49257664
KEYWORDS MGC.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Xenopus.
REFERENCE 1 (bases 1 to 2494)
AUTHORS Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W.
and Richardson,P.
TITLE Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative
JOURNAL Dev. Dyn. 225 (4), 384-391 (2002)
PUBMED 12454917
REFERENCE 2 (bases 1 to 2494)
AUTHORS Klausner,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,P.S., Wagner,L., Shennen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
```

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Scheezi, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Faney, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,  
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywinaki, M.I., Skalska, U., Smailus, D.E.,  
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932  
3 (bases 1 to 2494)  
Klein, S. and Gerhard, D.S.  
Direct Submission  
Submitted (23-JUN-2004) National Institutes of Health, Xenopus Gene  
Collection (XGC), National Institute of Child Health and Human  
Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD  
20892-7510, USA  
NIH-MGC Project  
Contact: XGC help desk  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement:  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
info@bcgsc.bc.ca  
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,  
Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth  
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,  
Kim MacDonald, Anara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,  
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabh, U,  
Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska,  
Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacques  
Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 161 Row: a Column: 22  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Hexamer frequency ORF  
analysis, Similarity but not identity to protein.

FEATURES  
source

1. 2494  
/organism="Xenopus laevis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8355"  
/clone="MGC:84197 IMAGE:6954181"  
/tissue\_type="Brain, adult xenopus"  
/clone\_lib="NICHD XGC\_Brn1"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6.1"  
1. 2494  
/gene="MGC84197"  
/db\_xref="LocusID:444656"  
86. 1804  
/gene="MGC84197"  
/codon\_start=1  
/product="MGC84197 protein"  
/protein\_id="AAH74350.1"  
/db\_xref="GI:49257665"  
/translaton="MESAHSTSLAVLCSVLSSVKRQSCVETLDPGRSLLSWRLRPRHG  
RESRPGICLTVLHQTGRLVNRTPQPSGGVTPVTVSVSGEATIDEXYNSMKW  
OHICKPAPFTVYQVRKHWRCKEYTFMCSDEQLFYQWLQAFHLLLEQOHTRPKNL  
LVYINPYGKRGKQIYENKVPALFSAAGTCAIVITEYANHARDHLYDANLEKYDGV  
VCVGGDGMFSEVLHGLVIRMQKDSVDHNNPNSAQLSRKNMRKIGIIPAGSTDCICATYAT

## gene

## CDS

GINDPETSAIHILILGDCQPLDVCSVHYKRTFLKYVSLLGYGFYDLVKGSEKNRWLG  
PARDVSGFTFLPHHCYEGSVSPQAKWVLGSPRODTCTTSGCYICRQSSKOLDEOE  
QTQACSEHREDDWTIIKGRFWAINAVSMSCACPRSPNGLSPAALHAGSDADLIIV  
RKSRDLPLHLIHTSKQDPFPFVEYVRVKNPQFTPKHPEDEDNESDITGLGKK  
NFSQICDHFSCGCHNVNSIWNCDGETLQDTAEMRVHQCULIKUFARGIENAKLEDV  
NPSKVF"

## ORIGIN

Alignment Scores:  
Pred. No.: 6,16e-109 Length: 2494  
Score: 1714.00 Matches: 322  
Percent Similarity: 72.32% Conservative: 83  
Best Local Similarity: 57.50% Mismatches: 125  
Query Match: 56.66% Indels: 30  
DB: 5 Gaps: 4

US-10-631-958-11 (1-562) x BC074350 (1-2494)

Qy	23	AlaAlaGluMetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLys	42
Db	83	GCCTGGAGTCCGCTCATTCACCTCGTTAGCCGCTGCTGCTCGGTCTCTCGGTTAAG	142
Qy	43	GlnGlnArgCysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSer	62
Db	143	AGGCAAAAGCTGCGAGGTGACCTGGACCCCGGGCGCTCTCTGCTCTCC---TGGAGAGAA	199
Qy	63	ProGlyProGlyAlaGly-----AlaProGly-----	71
Db	200	CTGAGGCCCGACACGGAAGAGAGAGCGGCGCGCGGAATCTCTTGACNAGAGTGCTC	259
Qy	72	-----AlaAspAlaCysSer	76
Db	260	CACCAAAACAGAGCCGTAGTAAACCGTTTCACTCGGCAATTCCTTCAGGTGGTGTACC	319
Qy	77	ValProValSerGluIleAlaValGluGluThrAspValHisGlyLysHisGlnGly	96
Db	320	GTCCAGTGAAGTAAATTTGTTTCAGTAGGGGAAGCAGAGATTGATGAGAGTACTACAAC	379
Qy	97	SerGlyLysTrpGlnLysMetGluLysProTrpAlaPheThrValHisCysValLysArg	116
Db	380	AGTAGAATGGCAGCATATTTGTAAGCCTCGTCGATTTACAGTACGTTATGTACAGAGA	439
Qy	117	AlaArgArgHisArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeu	136
Db	440	GTGAGGAAGCACCGCTGCGGTGCAAGGAGGTGACCTTTTGTGTCAGTGTAGCAGCTT	499
Qy	137	CysHisLeuTrpLeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLys	156
Db	500	TTTTATCAGTGGTTACAGCAATTTTCATGATCTACTTGGAGCAAACTCATAGACCAAG	559
Qy	157	HisLeuLeuValPheIleAsnProPheGlyLysGlyGlnGlyLysArgIleTrpGlu	176
Db	560	AACITGTAGTTTATATTAATCCTATGTGTGGGAAGAAAGAGGTAAACAGATATATGAG	619
Qy	177	ArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGlu	196
Db	620	AATAAAGTGGCTCCATTATTCAGTGTCTGGGATTTGTGCTGATGTCATAGTTACTGAG	679
Qy	197	HisAlaAsnGlnAlaLysGluThrLeuTrpGluIleAsnIleAspLysTrpAspGlyIle	216
Db	680	TATGCAAAACCATGTAGAGTACCTATATATGCCAACCTGGAAAAATATGACGGGGTT	739
Qy	217	ValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThr	236
Db	740	GTTTGTGTGGGGATGGAAATGTTTCAGCGAAGTGTGTCATGGCCTCATTTGTCAGATG	799
Qy	237	GlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLys	256
Db	800	CAAAAAGATCTGTATGATAGATACAAACAATCCAGTGCACAGTTATCTCGGTGTATATG	859
Qy	257	ArgIleGlyIleIleProAlaGlySerThrAspCysValCysTrpSerThrValGlyThr	276
Db	860	AGGATTGCAATTATCCCTGCTGGTTCAACAGACTGTATATGCTATGCTACTTGGCATT	919

Qy	277	SerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspVal	296
Db	920	AATGACCACAGAGAGCTCTGTTTGCACATCATCTCGTGACTGCCACGCTCTTGTATGTA	979
Qy	297	SerSerValHisAsnSerThrLeuLeuArgTrpSerValSerLeuLeuGlyTrpGly	316
Db	980	TGCTCTGTACACTACAAGAGGACATTTCTTAAGTATTTCGGTTTCGTTATTGGGATATGGT	1039
Qy	317	PheTrpGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTrp	336
Db	1040	TTTTATGAGATGTGTTAAAGGGCAGTGAAATAAATCGTTGGTTAGTCTCTGCTAGATAT	1099
Qy	337	AspPheSerGlyLeuLysThrPheLeuSerHisHisCysTrpGluGlyThrValSerPhe	356
Db	1100	GACGTTTCAGGTTCAAGACATTTCTGACTCACCATTGCTATGAAGGGTCAGTCTCATTC	1159
Qy	357	LeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCys	376
Db	1160	CAGCCAGCAAAAGTGGGTTCTGGGATCACCAGAGACCAACAACACTTGCACATCTGGGTGC	1219
Qy	377	PheValCysArgGlnSerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTrpGly	396
Db	1220	TACATTTGCGAGGCAAAATAGCAGCAGCTAGACGAGCAAGCAAGCAAAACCCAGGCTTGTGA	1279
Qy	397	LeuGluAlaAlaGluAspValGluTrpGlnValValCysGlyLysPheLeuAlaIle	416
Db	1280	TCAGAACACAGAGAACAGACGATGCTGGACACAATCAAGGAGATTCATGGCAATC	1339
Qy	417	AsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAla	436
Db	1340	AATGCTGTACGATGAGTGTGTCATGTCCAGAGCCCAATGGCCTTTCACACGACGCT	1399
Qy	437	HisLeuGlyAspGlySerSerAspLeuIleLeuArgLysCysSerArgPheAsnPhe	456
Db	1400	CACCTGGCAGATGGAAGTGCAGATCTCATTTAGTTAGAAAGTGCTCAAGACTTGAATTT	1459
Qy	457	LeuArgPheLeuIleArgHisThrAsnGlnAspGlnPheAspPheThrPheValGlu	476
Db	1460	TTAAGGCACCTTATCAGGCACACAAGCAATAAAGACCAGTTTGACTTTTCCATTTGTTGAA	1519
Qy	477	ValTrpArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAsp-----	494
Db	1520	GTTTATCGAGTAAAGAAATTTTCACTTACTCCAAAACACTTTGAAGATGAAGACAATGAA	1579
Qy	495	---SerAspLeuLysGluGlyLysLysArgPheGlyHisIleCysSerSerHisPro	513
Db	1580	AGCACAGATATAGAAATTTCTTGGCAAGAAAACCTTTTCTCAGATATGCACAGACCATCCA	1639
Qy	514	SerCysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSer	533
Db	1640	TCATGTGGATGAACCATGTGAATAGCATTTTGGAAATTTGTATGGGAGACATTGGACCAG	1699
Qy	534	ProAlaIleGluValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGlu	553
Db	1700	ACCGCCATTGAAATGAGGGTCCCACTGCCAACTATAAAATTAATTTGCACGCGGCATTGAG	1759

Search completed: September 6, 2005, 16:40:11  
Job time : 7486.62 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 03:56:01 ; Search time 30.7621 Seconds  
(without alignments)  
1363.781 Million cell updates/sec

Title: US-10-631-958-11  
Perfect score: 3025  
Sequence: 1 HEANGAPICVRAPPART.....QLVRLFARGIENPKPDHS 562

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pap:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pap:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pap:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pap:\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS\_COMB.pap:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	801.5	26.5	687	4	US-09-270-767-45874
2	585.5	19.4	359	4	US-09-270-767-46720
3	306	10.1	490	4	US-10-053-510-19
4	302.5	10.0	299	4	US-09-270-767-61405
5	301	10.0	524	4	US-10-053-510-20
6	271.5	9.0	618	4	US-09-970-516-4
7	271.5	9.0	618	4	US-09-817-676A-14
8	267.5	8.8	617	4	US-09-817-676A-12
9	266.5	8.8	384	4	US-09-959-897-2
10	264.5	8.7	384	4	US-09-970-516-2
11	263.5	8.7	384	4	US-09-949-016-7026
12	263.5	8.7	384	4	US-09-796-487-3
13	260	8.6	388	4	US-09-817-676A-15
14	260	8.6	388	4	US-09-796-487-2
15	255.5	8.4	368	4	US-10-053-510-21
16	254	8.4	373	4	US-09-796-487-5
17	254	8.4	381	4	US-09-796-487-1
18	254	8.4	381	4	US-09-796-487-4
19	247.5	8.2	382	4	US-09-970-516-6
20	244.5	8.1	392	4	US-09-796-487-6
21	242	8.0	424	4	US-09-796-487-8
22	228	7.5	536	4	US-09-248-796A-15859
23	206.5	6.8	204	4	US-09-796-487-9
24	198	6.5	403	4	US-09-796-487-7
25	161.5	5.3	312	4	US-09-949-016-9811
26	153	5.1	313	4	US-09-107-532A-7154
27	148	4.9	119	4	US-09-205-258-788

ALIGNMENTS

RESULT 1

US-09-270-767-45874  
; Sequence 45874, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270.767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 45874  
; LENGTH: 687  
; TYPE: PRT  
; ORGANISM: *Drosophila melanogaster*  
US-09-270-767-45874

Query Match 26.5%; Score 801.5; DB 4; Length 687;  
Best Local Similarity 31.3%; Pred. No. 9.7e-74;  
Matches 197; Conservative 97; Mismatches 206; Indels 129; Gaps 16;

Qy	36	QSVLWVQQRCAVSL	--PARALLRWRS	PGGAGACADACSV	PSVSEIIIAVEETDVHGK	93
Db	59	QQLVWERLQIKQSPQNEAKPL	----	PPDSAPQGGICSYGQSHVLHD	--DVVSI	111
Qy	94	HQSGKWKQMEKP	-----	YFTVHCVKRARR	-----HRWKWAQ	126
Db	112	RSQDTKASSLKPPSPGSRSSGCGSDVAQ	OGKPTSQYLTTINYAMRLSKSQ	QDTCNRWELRR	171	
Qy	127	VTFWCPEQLCHLWLTREMLEKT	---SRPKLLVFINPPGKGKRIYERKVA	PLF	183	
Db	172	LTFFNSDPYIVRWQDQLRLHSSSPTRMRVRL	LLFINPYGGRKGAQYERHVRIF	231		
Qy	184	TLASITTDIIITEHANQAKETLYEINDIKVD	IGVGVGDMFSVHLGLIGRTORSAGVD	243		
Db	232	QLAGVDATCITTORANQVKDILLSHDLGV	YDACCVCVGGDGTVAEVINGLI	FRQRELGLD	291	
Qy	244	QNHPRVLVPSLRIGIIPAGSTDCVCTYST	VTGTSALHIVVGSOLAMDVSSVHNS	303		
Db	292	EORPPYIPRP-ALPVGVIPAGSTDTIAYSM	HGTADVTAATHVLGHQHRGLDVCVSNQ	350		
Qy	304	TLLRYSVLLGYGYGDI	IKDSEKRWLGLARYDFSGLKTFLSHHCY	EGTVSFL	357	
Db	351	SLLFRCASVLSYGLGVAAQSENENRWMP	RRYISYGVKAFNLNRGYDAELRMLEEDLL	410		
Qy	358	---PAQTVGSPRD	-----RKPCRAGCFC	-----	RQSKQOLEEQKALYGL	398
Db	411	LTTPLEDIPQSPDSVCSLGSVP	SPVCYANCORCSFASSIQEORSLSLFIQESKEA	----E	466	

Sequence 264, App  
Sequence 38102, A  
Sequence 53319, A  
Sequence 328, App  
Sequence 5366, App  
Sequence 373, App  
Sequence 5133, App  
Sequence 62339, A  
Sequence 3740, App  
Sequence 3121, App  
Sequence 3813, App  
Sequence 4774, App  
Sequence 6786, App  
Sequence 45, Appl  
Sequence 37, Appl  
Sequence 1350, App  
Sequence 26145, A  
Sequence 7, Appl

```
Qy 399 AAEDVE-----EQVVGCKFLAINAT 419
Db 467 RNQVETEDSHLAASEAALLRPRPGNWLRLPTGSISSMRNLGNDQKVVGRGNFFMICGA 526
Qy 420 NMSCACRRSPGLSPAHLGCGSSDLILIRKSRFNFLRLIRHTNQO-DQDFDTFVEVY 478
Db 527 NITCACARSPNGISRYSHLGGCLDLILVKKTSLLNNVRLFLNTPAGRSGDIRNLPPFEVY 586
Qy 479 RVKKFOFTSKHMEEDSDLKGGKKRFCHIC-----SSHPSCCCTVSSNWCNCDGE 529
Db 587 RTRFRFRFASSEDYSLAGS-----CQPTTPEMTAHS-----STEFSSWNCNCDGE 635
Qy 530 VLSHSPAIEVRVHCQLVRLFARGIENPKP 558
Db 636 VVTDLDTMRSHCQLIEVFRGPHYSYKP 664

RESULT 2
US-09-270-767-46720
; Sequence 46720, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 46720
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46720

Query Match 19.4%; Score 585.5; DB 4; Length 359;
Best Local Similarity 35.5%; Pred. No. 8.7e-52;
Matches 125; Conservative 55; Mismatches 103; Indels 69; Gaps 6;

Qy 154 RPKHLLFVNPFGCGKGKRIYERKVAFLFTLASITTDIIVTEHANOAKETIYEINDKY 213
Db 13 RVRELLFVNPFGKKGAGQYRHRVPFIPOLAGVDATCTTQRANQVKDILLSHDIGVY 72
Qy 214 DGIVCVGGDGFSEVLHGLIGRTORSAGVDQNHPRVAVLVPSSLRIGIIPAGSTDVCYST 273
Db 73 DAVCCVGGDGTVAEVINGLIFQRMRELGLDQRPPYIPRP-ALPVGVIIPAGSTDTIAYSM 131
Qy 274 VGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGYPGYDIIKDSEKKRWGL 333
Db 132 HGTADVRTAAITHVLGQHRGLDVCVSGNQSLLRFCASVLSYGLGDVAAQSENRYRWGP 191
Qy 334 ARYDFGLKTLFSLHCHYEYGVFL-----PAQHTVGSPRD-----RKPCRAGC 376
Db 192 RRYEYGVKAPLNNRGYDAELRMLEEDPLLTLTLEDIPQSDVCSGLGSEVSPVCYANC 251
Qy 377 FVC-----ROSKQOLEEOKKALYGLEAAEDVE-----ERNQVETEDSHLAASEAALLRPRPGNLR 404
Db 252 QRCFSASIQORSSLFIOESKEA-----ERNQVETEDSHLAASEAALLRPRPGNLR 307
Qy 405 -----EQVVGCKFLAINATNMSCACRRSPRGLSPAHLGCG 441
Db 308 LPTGSISSMRNLGNDQKVVGRGNFFMICGANITCACARSPNGISRYSHLGDG 359

RESULT 3
US-10-053-510-19
; Sequence 19, Application US/10053510
; Patent No. 6830881
; GENERAL INFORMATION:
; APPLICANT: Saba, Julie D.
; APPLICANT: Fyset, Henrik
; TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
; ; TITLE OF INVENTION: POLYNUCLEOTIDES AND MODULATING AGENTS AND
```

```
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 200116.402C2
; CURRENT APPLICATION NUMBER: US/10/053,510
; CURRENT FILING DATE: 2002-01-17
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-053-510-19

Query Match 10.1%; Score 306; DB 4; Length 490;
Best Local Similarity 24.6%; Pred. No. 1.4e-22;
Matches 119; Conservative 68; Mismatches 165; Indels 132; Gaps 19;

Qy 135 QLCHLMQTLREMLE-KLTSRP-----KHLVFNPPGGKQGGKRIYERKVAFLFTLASIT 189
Db 12 QTAEMMHHTIRKHKRGNGSSPADCGKQLLILNPKSGSGKRELFOKVAPLLTEAEVQ 71
Qy 190 TDIIVTEHANOAKE--TLYEINIDYGVICVGGDMFSEVLHGLIGRTORSAGVDQNH 246
Db 72 YDIQITTHPOYAKEFVTRRDL-LTRYGLIVASGDGLFYEVNLGLMER-----MDW 122
Qy 247 PRAVLVPSSLRIGIIPAGSTDVCV-----CYSTVGTSDAETSALHIVVGDLSAMDVSSV 299
Db 123 RRAC---RELPLGIIPCGSGNGLAKSVAHHCNEPYEPKPIHLHATLTCMACKSTPMDVVRV 179
Qy 300 H---HNSTLLRYSVSLGYPGYDIIKDSEKKRWGLARYDFGLKTLFSLHCHYEYGVF 356
Db 180 ELATRDKHFVMSYFLSVGWLIIADIDIESERLSISGAQRFLLWAIRKLIGLRSYKGRVSY 239
Qy 357 -----LPAQHTVGSPRDRKPCRAGCFVCRQSKQOLEEOKKALYGLEAA 400
Db 240 LLGKGGKKEPPVEAARELPRAESTAGIRSSLPINAGEF---HDLPEEEREAVIDGEQFA 295
Qy 401 EDV-----EEQVVC-----GKFLAINATNMSCACR-----RSP 429
Db 296 DAISLDRSVVRQHADSWHSAMSRRTAYISLGGPSMRNSRMSISQRIEANAEEFAERVP 355
Qy 430 RG-----LSPAHLGCGSSDLILIRK- 450
Db 356 TGTIPPLQMPLLSSDGCWICEDGDFVMVHAAATHTLSSDVFVFAESRLDDGLIYLVIIRRG 415
Qy 451 CSRFNFLRLIRHTNQDQF---DFTFVEYVRVKKFOFTSKHMEDESD---LKEGGKKR 504
Db 416 VSRHQLLNFL-NLNAGTHLPICGEDPFIVKVPCEAFR-----IEPSSSDGILLVVDGERVE 469
Qy 505 FGH 508
Db 470 YGPI 473

RESULT 4
US-09-270-767-61405
; Sequence 61405, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61405
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-61405

Query Match 10.0%; Score 302.5; DB 4; Length 299;
Best Local Similarity 29.2%; Pred. No. 1.5e-22;
```



Matches 85; Conservative 35; Mismatches 82; Indels 89; Gaps 9;

QY 342 KTFLEHHCVEGVTSFL-----PAQHTVGSPPD-----RKPCRAGCFVC----- 379  
DB 1 KAFLLNNGYDAELRMLEEDPDLITTPLEDIPQSPDSVCSGLGSVSVYANQRCSPASS 60  
QY 380 ---ROSKQOLEEKALYGLEAEDVE----- 404  
DB 61 IQEQSSSLFIQESKEA-----ERNQVETEDSHLAASEALLRPRPRGNLRLPTGSSISS 116  
QY 405 -----EMQVCGKFLAINATNMSCARRSPRGLSPAHLGDGSSDLILIRKCSRNFPL 457  
DB 117 MENLNDQWKVVRGNFFMCGANITCACARSPNGISRYSHLGDGCLDLILVKKTSLLNNV 176  
QY 458 RFLIRHTNQ-QDPDFTFVEVYRVKFTSKHMEDESDSKKEGKKRFGHC----- 509  
DB 177 RFLNTAGSGDIRLPLFVEVYTRFRTRTSASEEDYSLAGS-----CQPIPTPE 228  
QY 510 ---SSHPSCCTVSNSSWNCDCGVLHSPALFVVRVHCOLVRLFARGIEENPKP 558  
DB 229 ENTAAHS---STEFSSWNCDEGVVTDLDLTKRSHCOLIEVFMRGPHSYSKP 276

RESULT 5  
US-10-053-510-20  
; Sequence 20, Application US/10053510  
; Patent No. 630881  
; GENERAL INFORMATION:  
; APPLICANT: Saba, Julie D.  
; APPLICANT: Fyrist, Henrik  
; TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND MODULATING AGENTS AND  
; TITLE OF INVENTION: METHODS OF USE THEREFOR  
; FILE REFERENCE: 200116.402C2  
; CURRENT APPLICATION NUMBER: US/10/053,510  
; CURRENT FILING DATE: 2002-01-17  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 20  
; LENGTH: 524  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-10-053-510-20

Query Match 10.0%; Score 301; DB 4; Length 524;  
Best Local Similarity 25.2%; Pred. No. 5e-22;  
Matches 121; Conservative 66; Mismatches 151; Indels 142; Gaps 20;

QY 140 WLQTLREMLEK-----LTSRPKLLVFINPFGGKGKRIYERKVAFLPLASIT 189  
DB 17 MYRSLRWQLHRTLEEIFVAPTVDERRRVLNLPKSGSGDAREVFNNMHTVPLNEAEVP 76  
QY 190 TDIIIVTEHANOAKETLYEINIDKYGIVCVGGDMFSEVLHGLIGRTORSAGVDQNHPR 249  
DB 77 YDLYTKHNSFAIEFLSTRCLDAWCCVAVAGDGLFHEIVNGLLRQ-----DWAH--- 127  
QY 250 VLVPSLSLRIGIIPAGSTDCV-----CYS--TVGTSDAETSALHIVGDSLAMDVSSVHH 301  
DB 128 --VLPHLALGIIPCGSGNGLARSIAHCYNKPVLG-----AALTVISGRSSPMDVVRVOL 179  
QY 302 NSTLRYSVLLGYGFYGDIDKSEKRWGLARYDFSGLKTFLSHHCYEGTVSFLPAQH 361  
DB 180 QSRSL-YSPFLSIGWGLISDVDIESIRMLGYQRTVMTLYRLVNLRTYNGRISYLLTDH 238  
QY 362 TVGSPRD-----RKPCRACFP-----VCRQSKQOLEE----- 389  
DB 239 EVSSTHSATGYAQRMRQSSRCNTHIDMLNGPAPIYHSSAEYLPOEFADVISLTSINQ 298  
QY 390 -----QKALY-----GLEAEE-----DV 403  
DB 299 SFRSCDSWLSGSRRSFYYSISEIYHSLADESEFAGLAASLENRQNYGPASELPDL 358  
QY 404 EE-----WQVVCCKFLAINA---TNMSCACRRSPRGLSPAHLGDGSSDLILIRK-CS 452

DB 359 NEPLSEDOGWLVEGEFVMMHAYVQTHLGDICH-----FAPKALNDGTIYLLIRAGIS 413  
QY 453 RFNFLRFLR-----HTNQDQDFDFTFVEVYRVKFTSKHMEDESD---LKEGKGKRF 506  
DB 414 RPHLLSFLYNNSSGTHLPESH-DHVKVLPVRAFR-----LEPYDNHGIIITVDGERVEFG 467

RESULT 6  
US-09-970-516-4  
; Sequence 4, Application US/09970516  
; Patent No. 6610534  
; GENERAL INFORMATION:  
; APPLICANT: No. 6610534artis AG  
; TITLE OF INVENTION: Induction of blood vessel formation through administration of  
; FILE REFERENCE: 4-31617  
; CURRENT APPLICATION NUMBER: US/09/970,516  
; CURRENT FILING DATE: 2001-10-04  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 4  
; LENGTH: 618  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-970-516-4

Query Match 9.0%; Score 271.5; DB 4; Length 618;  
Best Local Similarity 30.0%; Pred. No. 7.5e-19;  
Matches 78; Conservative 44; Mismatches 111; Indels 27; Gaps 4;

QY 116 RARHRKWAQVTPWCPPEEQCLHWLQTLRMLSEKLTSPKLLVFINPFGKGKGRYI 175  
DB 111 RAEQRWATAITCL-----LRGLPFGDGIITPDLPRPRLLLLVNPFQGLAWQMC 164  
QY 176 ERKVAFLPLASITTDIIIVTEHANOAKETLYEINIDKYGIVCVGGDMFSEVLHGLIGR 235  
DB 165 KHVLPMLISEAGLSFNLIQTERQHARELVGGLSEWDGIVTVSGDGLLHEVLNGLLDR 224  
QY 236 TORSAGVDQNHPRAVLPVPSLSLRIGIIPAGSTDCVCYS-----TVGTSDAETSALHI 286  
DB 225 -----PDWEEAVKMP-----VGLPCGSGNALAGAVNQHGFPALGLDLLNCSLL 272  
QY 287 VVGDSLAMDVSSVHHNSTLLRYSVLLGYGFYGDIDKSEKRWGLARYDFSGLKTFLS 346  
DB 273 CRGGHPDLDLISVTLASGRCFSLSVANGFVSDVIQSERFALGSARFTLTGLTVLGLAT 332  
QY 347 HHCYEGTVSFLPAQHTVGVSP 366  
DB 333 LHTYGRSLYLPATVEPASP 352

RESULT 7  
US-09-817-676A-14  
; Sequence 14, Application US/09817676A  
; Patent No. 6800470  
; GENERAL INFORMATION:  
; APPLICANT: Spiegel, Sarah  
; APPLICANT: Kohama, Takafumi  
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,  
; TITLE OF INVENTION: Expression and Methods of Use Thereof  
; FILE REFERENCE: 00170/HG  
; CURRENT APPLICATION NUMBER: US/09/817,676A  
; CURRENT FILING DATE: 2001-03-26  
; PRIOR APPLICATION NUMBER: US 60/194,318  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patent in ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 618  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-817-676A-14



Query Match	9.0%	Score	271.5;	DB	4;	Length	618;
Best local Similarity	30.0%;	Pred.	No. 7.5e-19;				
Matches	78;	Conservative	44;	Mismatches	111;	Indels	27;
Gaps	4;						
Qy	116	RARRHRKWAQVTCWCP	EEOLCHLWQLT	LRMLEKLSRPHKL	LVFINPFGG	GKGKRIY	175
Db	111	RAEQRWATALTCL	-----LRGLP	PDGGEITPDLL	PRPRLULLVNP	FGGRGLANQWC	164
Qy	176	ERKVAPLFTLASITTD	IIVTEHANOAKET	LVINIDKYD	GVCGDMFSEVLH	GLIGR	235
Db	165	KNHVLPMISEAGLSFN	LQIOTERQHAREL	VQGLSEWDG	IVTVSGDGLLH	EVNLGLLDR	224
Qy	236	TQRSAGVDQNHPRAV	LPVSSLRIGII	PAGSTD	CVCS-----	TVGSDAETSALHI	286
Db	225	-----POMEAAVMP	---VGILPCGS	NALAGAVNQHG	GFEPALGDL	LLNCULL	272
Qy	287	VVGDSLAMDYSSVHH	NSTLLRYSVSL	GYGYGII	IKDSEKKRWGL	GLARYDFSL	346
Db	273	CRGGCHPLDLLSVTL	ASGRCSFSLSV	ANGVPSVD	IQSEFRALG	SARFTLGT	332
Qy	347	HHCEGTVSFLPAO	HTVGPSP	366			
Db	333	LHTYGRGLSYLPAT	VEPASP	352			

RESULT 8  
 US-09-817-676A-12  
 ; Sequence 12, Application US/09817676A  
 ; Patent No. 6800470  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Spiegel, Sarah  
 ; APPLICANT: Kohama, Takafumi  
 ; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,  
 ; TITLE OF INVENTION: Expression and Methods of Use Thereof  
 ; FILE REFERENCE: 00170/HG  
 ; CURRENT APPLICATION NUMBER: US/09/817.676A  
 ; CURRENT FILING DATE: 2001-03-26  
 ; PRIOR APPLICATION NUMBER: US 60/194,318  
 ; PRIOR FILING DATE: 2000-04-03  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 12  
 ; LENGTH: 617  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 US-09-817-676A-12

```

RESULT 9
US-09-959-897-2
: Sequence 2, Application US/09599897
: Patent No. 6730480
: GENERAL INFORMATION:
: APPLICANT: PITSON, Stuart M
: APPLICANT: Brian, WATTENBERG W
: APPLICANT: Pu, XIA
: APPLICANT: Richard, D'ANDREA J
: APPLICANT: Jennifer, BAMBLE R
: APPLICANT: Mathew, VADAS A
: TITLE OF INVENTION: SPHINGOSINE KINASE ENZYME
: FILE REFERENCE: PITSON=1
: CURRENT APPLICATION NUMBER: US/09/959,897
: CURRENT FILING DATE: 2001-11-13
: PRIOR APPLICATION NUMBER: PCT/AU00/00457
: PRIOR FILING DATE: 2000-05-12
: PRIOR APPLICATION NUMBER: AU PQ 0339
: PRIOR FILING DATE: 1999-05-13
: PRIOR APPLICATION NUMBER: AU PQ 1504
: PRIOR FILING DATE: 1999-07-08
: NUMBER OF SEQ ID NOS: 56
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 2
: LENGTH: 384
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-959-897-2

```

RESULT 10  
US-09-970-516-2  
; Sequence 2, Application US/09970516  
; Patent No. 6610534  
; GENERAL INFORMATION:  
; APPLICANT: NO. 6610534artis AG  
; TITLE OF INVENTION: Induction of b

; TITLE OF INVENTION: polynucleotides encoding sphingosine kinases

; FILE REFERENCE: 4-31617

; CURRENT APPLICATION NUMBER: US/09/970,516

; CURRENT FILING DATE: 2001-10-04

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn version 3.11

; SEQ ID NO 2

; LENGTH: 384

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-970-516-2

Query Match 8.7%; Score 264.5; DB 4; Length 384;

Best Local Similarity 25.3%; Pred. No. 1.9e-18;

Matches 107; Conservative 68; Mismatches 169; Indels 79; Gaps 16;

Qy 154 RPKHLLVFINPFGGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLYEINIDKY 213

Db 13 RPCRVLVLLNPRGGKGKALQFRSHVQPLLAEAESFTLMLTERNHARELVRSEELGRW 72

Qy 214 DGIVCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGSTDCVC--- 270

Db 73 DALVVMGSDGLMHVEVVGMLMERPDWETAIQK-----PLCSLPAGSGNALAASL 120

Qy 271 -----YSTVGTSDAETSALHIVVGDLSA-MDVSSVHHNSTLLRYSVSLGFGFYGDIIKD 324

Db 121 NHYAGYEQVTNEDLLTNCTLLCRLLSPMNLLSLHTASGLRFLSVLSLAWGFTADVLE 180

Qy 325 SEKKRWLGLARYDFSLGTLFSLSHCYEGTVSFLPAQHTVGSPPDRKPCRAGCFVCRQSKQ 384

Db 181 SEKYRLGEMRFTLTGTLRLAALRYRGLAYLPVGR-VGSKTPASP-----VVVQGPV 234

Qy 385 QLEEEQKALYGLEAAEDV-BEWOVCGK-----FLAINATNMSCACRRSPRGLSPAHLG 439

Db 235 DAH-----LVPLE--EPVPSHWTVPDDEDFVLVLLALLSHLGSEMFAPMGRCAA----- 282

Qy 440 DGSSDLLILRK-CSRFNPLRLI-----RHTNQOQPDFTFVYVYRVKKFQFTSKHMEDESD 497

Db 283 -GVMHLLFYVRAGVSRAMLLRFLAMEKGRHM-----EYECPLYLVVVPVAVFR-----LBPK--- 333

Qy 498 KEGGKRFKHCSSHPSCCTVSNSSMNCDEVLHSPAIEVRVHCQLVRLFARGIEENP- 556

Db 334 ---GKGVFA-----VDGELMVSEAVQGVHPNYFMVMSGCVPEPPS 371

Qy 557 -KP 558

Db 372 WKP 374

#### RESULT 11

US-09-949-016-7026

; Sequence 7026, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7026

; LENGTH: 384

; TYPE: PRT

; ORGANISM: Human

US-09-949-016-7026

Query Match 8.7%; Score 263.5; DB 4; Length 384;

Best Local Similarity 25.1%; Pred. No. 2.4e-18;

Matches 107; Conservative 69; Mismatches 164; Indels 87; Gaps 17;

Qy 154 RPKHLLVFINPFGGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLYEINIDKY 213

Db 13 RPCRVLVLLNPRGGKGKALQFRSHVQPLLAEAISFTLMLTERNHARELVRSEELGRW 72

Qy 214 DGIVCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGSTDCVC--- 270

Db 73 DALVVMGSDGLMHVEVVGMLMERPDWETAIQK-----PLCSLPAGSGNALAASL 120

Qy 271 -----YSTVGTSDAETSALHIVVGDLSA-MDVSSVHHNSTLLRYSVSLGFGFYGDIIKD 324

Db 121 NHYAGYEQVTNEDLLTNCTLLCRLLSPMNLLSLHTASGLRFLSVLSLAWGFTADVLE 180

Qy 325 SEKKRWLGLARYDFSLGTLFSLSHCYEGTVSFLPAQHTVGSPPDRKPCRAGCFVCRQSKQ 384

Db 181 SEKYRLGEMRFTLTGTLRLAALRYRGLAYLPVGR-VGSKTPASP-----VVVQGPV 234

Qy 385 QLEEEQKALYGLEAAEDV-BEWOVCGK-----FLAINATNMSCACRRSPRGLSPAHLG 439

Db 235 DAH-----LVPLE--EPVPSHWTVPDDEDFVLVLLALLSHLGSEMFAPMGRCAA----- 282

Qy 440 DGSSDLLILRK-CSRFNPLRLI-----RHTNQOQPDFTFVYVYRVKKFQFTSKHMEDE 493

Db 283 -GVMHLLFYVRAGVSRAMLLRFLAMEKGRHM-----EYECPLYLVVVPVAVFR-----LBPK 332

Qy 494 DSDLKEGKRFKHCSSHPSCCTVSNSSMNCDEVLHSPAIEVRVHCQLVRLFARGIE 553

Db 333 D-----GKGVFA-----VDGELMVSEAVQGVHPNYFMVMSGCV 367

Qy 554 ENP--KP 558

Db 368 PPSWKP 374

#### RESULT 12

US-09-796-487-3

; Sequence 3, Application US/09796487

; Patent No. 6830916

; GENERAL INFORMATION:

; APPLICANT: Spiegel, Sarah

; TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use

; FILE REFERENCE: 07320001aa (2033957-0001)

; CURRENT APPLICATION NUMBER: US/09/796.487

; CURRENT FILING DATE: 2001-03-02

; PRIOR APPLICATION NUMBER: US 60/186,532

; PRIOR FILING DATE: 2000-03-03

; PRIOR APPLICATION NUMBER: US 09/530,868

; PRIOR FILING DATE: 2000-05-05

; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3

; LENGTH: 384

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: MISC FEATURE

; LOCATION: (1)..(384)

; OTHER INFORMATION: SEQ ID NO 3 is the peptide sequence of hSPHK1 in Fig. 3, correspo

; nding to amino acid residue 1 to 384 of Homo sapiens SPHK-1 of Ge

; OTHER INFORMATION: nBank sequence Accession Number AAF73423.

; PUBLICATION INFORMATION:

; AUTHORS: Nava et al.

; TITLE: Functional characterization of human sphingosine kinase-1

; JOURNAL: FEBS Lett.

; VOLUME: 473

; ISSUE: 1

; PAGES: 81-84

; DATE: 2000

; DATABASE ACCESSION NUMBER: AAF73423

```
; DATABASE ENTRY DATE: 2000-06-01
; RELEVANT RESIDUES: (1)..(384)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AF73423
; DATABASE ENTRY DATE: 2000-06-01
; RELEVANT RESIDUES: (1)..(384)
US-09-796-487-3

Query Match      8.7%; Score 263.5; DB 4; Length 384;
Best Local Similarity 25.1%; Pred. No. 2.4e-18;
Matches 107; Conservative 69; Mismatches 164; Indels 87; Gaps 17;

154 RPKHLVFINPFGGKGKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLYEINIDKY 213
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 13 RPCRVLNLPFGGKGKALQFRSHVQPLLAEEISFTLMTERRNHARELVRSEELGRW 72

214 DGIYCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGSTDVCV--- 270
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 73 DALVMSGDGLMHVEVVGMLMERPDWETAIQK-----PLCSLPAGSGNALAASL 120

271 -----YSTVGTSDAETSALHIVVGDLSA-MDVSSVHNSTLLRYSVSLGFGYVDI1KD 324
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 NHYAGYEQVTNEDLLTNTCLLCRRLLSPMNLISHTASGLRFLSVLSLWGFADVDLE 180

325 SEKKRWGLARYDFSLGKTLFSLHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFYCQSKQ 384
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 SEKYRRLGEMRFTLTGTLRLAALRTYGRLAYLPVGR-VGSKTPASP-----VVVQGPV 234

385 QLEEQKALYGLAAEDV-EWQVVCVK-----FLATNATMSCACRRSPRGLSPAHLG 439
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 235 DAH-----LVPLR--EPVPSHMTVPDDEFLVILALLHSLGSEMPAAMPGRCAA---- 282

440 DGSSDLILIRK-CSRFNPLRLI-----RHTNQDQDFTFVEYVRVKKQFTSKHMEDE 493
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 283 -GVHMLFYVRAGVSRAMLLRFLAMEKGRM-----EYECPLVTVVPVAFR-----LEPK 332

494 DSDLKEGKRFHICSHSPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIE 553
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 333 D-----GKGVFA-----VDGELMVSEAVQGVHPNYFMVMVSGCVE 367

554 ENP--KP 558
   |||
Db 368 PPSWKP 374

RESULT 13
US-09-817-676A-15
; Sequence 15, Application US/09817676A
; Patent No. 6800470
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817,676A
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/194,318
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Mus musculus
; PUBLICATION INFORMATION:
; TITLE: Molecular cloning and functional characterization of
; TITLE: murine sphingosine kinase
; JOURNAL: J. Biol. Chem.
; VOLUME: 273
; ISSUE: 37
; PAGES: 23722-23728
; DATE: 1998-09-11
```

```
; DATABASE ACCESSION NUMBER: AAC61698
; DATABASE ENTRY DATE: 1998-09-26
US-09-817-676A-15

Query Match      8.6%; Score 260; DB 4; Length 388;
Best Local Similarity 23.3%; Pred. No. 5.6e-18;
Matches 100; Conservative 70; Mismatches 173; Indels 86; Gaps 13;

129 FWCPEEQCHLWLOTLREMLEKLTSPKHLVFINPFGGKGKRIYERKVAFLFTLASI 188
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2 WWC-----CVLFVV---ECPRGLLPFCRVVLINPOGGKGKALQLFQSRVQFLEAEI 53

189 TTDIIIVTEHANOAKETLYEINIDKYIGIVCVGGDGMFSEVLHGLIGRTORSAGVDQNHPR 248
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 54 TFXLILTERKNHARELVCABEELGHWDALAVMSGDGLMHVEVVGMLMERPDWETAIQK--- 109

249 AVLVPSSLRIGIIPAGSTDVCV-----YSTVGTSDAETSALHIVVGDLSA-MDVSSV 299
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 -----PLCSLPAGSGNALAASVNHVAGYEQVTNEDLLINCTLLCRRLSPMNLISL 161

300 HNSTLLRYSVSLGFGYVDI1KDSEKKRWGLARYDFSLGKTLFSLHHCVEGTVSFLPA 359
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 162 HTASGLRLYSVLSLWGFADVDLESEKYRRLREIRFTVGTFFRLASRLRYQGLAYLP- 220

360 QHTVGSPPDRKPCRCAGCFYCQSKQQLLEEQKALYGLAAEDV-EWQVVCVK-----FL 414
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 221 ---VGTVASKRPA-----STLVQKGPVDTHLVPLEEPVPVSHMTVPVPEQDFVLVL 266

415 AINATMSCACRRSPRGLSPAHLGDSSDLILIRK-CSRFNPLR-FLIRHTNQDQDFD 472
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 267 VLLHHTLSSSELFAAMPGRCEA-----GVHMLFYVRAGVSRAMLLRFLAMEKGMELDC 321

473 TFVEYVRVKKQFTSKHMEDESDLEKGGKRFHICSHSPSCCTVSNSSWNCDEVLH 532
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 322 PYLVHVPVAFRLEPRS-----QRGVFSVDGELMV 351

533 SPAIEVRVH 541
   |||:|||||
Db 352 CEAVQGVH 360

RESULT 14
US-09-796-487-2
; Sequence 2, Application US/09796487
; Patent No. 6830916
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
; FILE REFERENCE: 07320001aa (2033957-0001)
; CURRENT APPLICATION NUMBER: US/09/796,487
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,532
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: US 09/530,868
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Mus Musculus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)-(388)
; OTHER INFORMATION: SEQ ID NO 2 is the peptide sequence of SPHK1b in Fig. 1, correspo
; OTHER INFORMATION: nding to amino acid residue 1 to 388 of SPHK1b of GenBank sequenc
; OTHER INFORMATION: e Accession Number AAC61698.
; PUBLICATION INFORMATION:
; AUTHORS: Kohama et al.
; TITLE: Molecular cloning and functional characterization of murine sphingosine
; TITLE: kinase
; JOURNAL: Journal of Biological Chemistry
; VOLUME: 273
```

```

; ISSUE: 37
; PAGES: 23722-23728
; DATE: 1998
; DATABASE ACCESSION NUMBER: AAC61698
; DATABASE ENTRY DATE: 1998-09-26
; RELEVANT RESIDUES: (1) .. (388)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AAC61698
; DATABASE ENTRY DATE: 1998-09-26
; RELEVANT RESIDUES: (1) .. (388)
US-09-796-487-2

Query Match      8.6%; Score 260; DB 4; Length 388;
Best Local Similarity 23.3%; Pred No. 5,6e-18;
Matches 100; Conservative 70; Mismatches 173; Indels 86; Gaps 133

Qy 129 FWCPEEQCHLWLOTLREMLEKLTSPRPKHLVFINPFGKGQKRIYBRKVAPLFTLASI 188
Db 2 WWC-----CVLFVW---ECPRGLLPKPCRVLLNPQGGKGKALQLFQSRVQPFLEEAEI 53

Qy 189 TTDLIVTEHANOAKETLYEINIDKYDGIVCVGGDGMFSEVHLGLIGRTQRSAGVDQNHPR 248
Db 54 TFKULITERKQWARELVCABELGHDALAVMSGDGLMEHVNGLMERPDWETAQK----- 109

Qy 249 AVLVPSSLRIGIIPAGSTDVC-----YSTVTGTSDAETSALHVVVGDSLA-MDVSSV 299
Db 110 -----PLCSLPGGSGNALAASVNHVAGVEQVTNEDLLINCTLLCRRRLSPMNLJSL 161

Qy 300 HHNSTLURYSVLLGYGYPGDIIDKSEKKRMILGARYDFSGLKTFLSHHCYEGTVSFLPA 359
Db 162 HTASGLRLYSVLSWSGFVADVDLESEKRYRLGRIFTVGTFFRLASLRITYQGQLAYLP- 220

Qy 360 QHTVGSPRDRKPCFACGFCVCRSQKQLEEQEKALYGLEAAEDV--EEQVVCVKG-----FL 414
Db 221 ---VGTVASKEPA-----STLVQKGPVDTHLVLEPVPSHWTVVPEQDFVLVL 266

Qy 415 AINATNMSCACRRSPRGLSPAHLGDSGSDILIRK-CRRNFILR-FLIRHTNQOQPDF 472
Db 267 VLLHTHLSSELFAAPMGCEA-----GYMHLFYVRAGVSRALLRLFLAMQKGKHELDJC 321

Qy 473 TFVEYVYKVKFQFTSKHMEDESDLKKEGKKRFGHICSSHPSCCCTVSNSSWNCDCVLIH 532
Db 322 PYLVHVPVVAFLRPR-----OQGVFSDVDELNV 351

Qy 533 SPAIEVRVH 541
Db 352 CEAVQGOVH 360

```

```

RESULT 15
US-10-053-510-21
; Sequence 21, Application US/10053510
; Patent No. 6830881
; GENERAL INFORMATION:
; APPLICANT: Saba, Julie D.
; APPLICANT: Fyrest, Henrik
; TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
; TITLE OF INVENTION: POLYNUCLEOTIDES AND MODULATING AGENTS AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 200116.402C2
; CURRENT APPLICATION NUMBER: US/10/053.510
; CURRENT FILING DATE: 2002-01-17
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-053-510-21

```

Query Match	8.4%;	Score 255.5;	DB 4;	Length 368;
Best Local Similarity	25.1%;	Pred. No. 1.5e-17;		
Matches 105; Conservative	68;	Mismatches 167;	Indels 79;	Gaps 16;

Qy	158	LLVFNPFGKGQGR	IYERK	VAPLFT	LTASIT	TDII	IVTEH	ANQAKET	LYEIN	DKYD	GV	217
Db	1	VLVLNPRGGKG	KALQF	RSHVQ	PLLAAE	AEIS	FTL	MLTER	NHAREL	VRSEEL	GRDA	60
Qy	218	CVGGDGMFSE	VHLGLIG	TQRSAG	VQDNH	PRAVL	VPSS	LRIGI	IIPAG	STDC	VCV	270
Db	61	VMSGDGLM	HEVVG	MGMR	PDWET	AIQ	-----	PLCSL	PAGSG	NLAAS	LNHYA	108
Qy	271	-YSTVGTSD	AETSL	ALHIV	VVGDS	LA-MDV	SSVHH	NSTL	RYSV	SLG	GYFG	328
Db	109	GYEQVTNE	DLTNC	TLLCR	ELLSP	MNLL	SLHT	ASGL	RLFS	VLSL	AWGFI	168
Qy	329	RWLG	LARYD	FSGL	TKTF	SHHCY	EGTV	SVSL	PAQHT	VGS	PRDRK	388
Db	169	RLGEM	FTLGT	FLURL	AALRT	YRGL	ALPL	VGR-V	GSKT	PASP	-----	221
Qy	389	EQKAL	YGLEA	AEAD	V-EWQ	VVCK	---E	FLA	INAT	NMSC	ACRR	443
Db	222	-----	LVPLE	--EP	VP	SHW	TV	PD	EDF	VL	UAL	269
Qy	444	DLIL	IRK-C	SFR	NELR	-FL	IHT	NQ	QD	FD	TF	501
Db	270	HLFV	YRAG	VSEAM	LLRL	FLA	MKG	RHME	YEC	PYL	VYV	318
Qy	502	KGR	FGH	ICSS	HPSC	CC	TV	SNSS	WNC	DGE	VLH	558
Db	319	KG	VEA	-----	VDGE	LMY	SEAV	QVQ	GHV	PNY	FMM	358

Search completed: September 3, 2005, 04:10:38  
Job time : 32.7621 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 04:09:18 ; Search time 114.766 Seconds  
(without alignment)  
1893.930 Million cell updates/sec

Title: US-10-631-958-11  
Perfect score: 3025  
Sequence: 1 HEANGPAPLGVAPPWRT.....QLVRLFARGIENKPDSDS 562

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : A\_Geneseq\_16Dec04:\*
- 1: Geneseqp1980s:\*
  - 2: Geneseqp1990s:\*
  - 3: Geneseqp2000s:\*
  - 4: Geneseqp2001s:\*
  - 5: Geneseqp2002s:\*
  - 6: Geneseqp2003as:\*
  - 7: Geneseqp2003bs:\*
  - 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3025	100.0	562	ABB07857	Abb07857 Human sph
2	2888	95.5	537	ABB07856	Abb07856 Human sph
3	2888	95.5	537	ADA05680	Ada05680 Human NOV
4	2888	95.5	537	ADJ96664	Adj96664 Human lip
5	2888	95.5	537	ADN62845	Adn62845 Human NOV
6	2880	95.2	537	AAW49115	Aaw49115 Human cer
7	2588	85.6	481	ABR56302	AbR56302 Human Sph
8	2463	81.4	460	AAy96059	Aay96059 Human sph
9	2456.5	81.2	471	AAE07884	Aae07884 Human sph
10	2428	80.3	531	ADP55248	Adp55248 Human PRO
11	2350	77.7	536	ADs11054	Adsl1054 Human the
12	2215	73.2	746	ABG13541	Abg13541 Novel hum
13	2210	73.1	416	ABR56301	AbR56301 Human Sph
14	1714.5	56.7	727	ABG13543	Abg13543 Novel hum
15	1640.5	54.2	326	ABB07854	Abb07854 Human sph
16	1361.5	45.0	454	ABG13544	Abg13544 Novel hum
17	1159	38.3	255	ADs12267	Adsl2267 Human the
18	1055	34.9	228	AA841822	Aa841822 Human ORP
19	1032	34.1	190	ABP64913	Abp64913 Human pro
20	986	32.6	182	AA842383	Aa842383 Human ORF
21	675	22.3	596	ABB69669	Abb69669 Drosophil
22	664	22.0	136	ABG13540	Abg13540 Novel hum
23	647	21.4	144	AAE07885	Aae07885 Partial r
24	628	20.8	136	ABG13542	Abg13542 Novel hum
25	600	19.8	532	ADQ88891	Adq88891 Novel hum

ALIGNMENTS

RESULT 1  
ABB07857

ID ABB07857 standard; protein; 562 AA.

AC ABB07857;

DT 03-JUL-2002 (first entry)

DE Human sphingosine kinase-like protein.

KW Human sphingosine kinase-like protein; intracellular signalling;

KW cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;

KW autoimmune disease; rheumatoid arthritis; Parkinson's disease.

OS Homo sapiens.

PN WO200228906-A2.

PD 11-APR-2002.

PF 05-OCT-2001; 2001WO-EP011516.

PR 06-OCT-2000; 2000US-0238005P.

PR 23-AUG-2001; 2001US-0314113P.

XX (FARB ) BAYER AG.

XX Kossida S, Encinas J;

XX WPI; 2002-340094/37.

XX N-PSDB; ABL40828.

XX New reagent for modulating the activity of sphingosine kinase-like protein polypeptide or polynucleotide and treating cancer, asthma, allergy, an autoimmune disease, or a central or peripheral nervous system disorder.

XX Claim 25; Fig 11; 120pp; English.

XX The invention relates to a human sphingosine kinase-like protein. The polypeptide can be expressed by standard recombinant methodology. The sphingosine kinase-like protein and gene can be used to regulate intracellular signalling and consequently cell proliferation and apoptosis. Such regulation is useful for treating cancer, allergies (e.g. asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and peripheral nervous system disorders (e.g. Parkinson's disease). The present sequence represents the human sphingosine kinase-like protein





DE Human NOV9a protein SEQ ID NO:40.

XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;

KW immunomodulator; cytosatic; nootropic; neuroprotective;

KW antiparkinsonian; antilipaeic; gene therapy; human disease;

KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;

KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;

KW immune disorder; haematopoietic disorder; dyslipidaemia.

XX

OS Homo sapiens.

XX

XX WO2003029424-A2.

XX

XX 10-APR-2003.

XX

XX 02-OCT-2002; 2002WO-US031373.

XX

XX 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 05-OCT-2001; 2001US-0327449P.

PR 09-OCT-2001; 2001US-0327917P.

PR 09-OCT-2001; 2001US-0328029P.

PR 09-OCT-2001; 2001US-0328044P.

PR 09-OCT-2001; 2001US-0328056P.

PR 12-OCT-2001; 2001US-0328849P.

PR 15-OCT-2001; 2001US-0329414P.

PR 17-OCT-2001; 2001US-0330142P.

PR 18-OCT-2001; 2001US-0330309P.

PR 22-OCT-2001; 2001US-0341058P.

PR 24-OCT-2001; 2001US-0343629P.

PR 29-OCT-2001; 2001US-0349575P.

PR 01-NOV-2001; 2001US-0346357P.

PR 17-APR-2002; 2002US-0373260P.

PR 19-APR-2002; 2002US-0373815P.

PR 19-APR-2002; 2002US-0373817P.

PR 19-APR-2002; 2002US-0373826P.

PR 19-APR-2002; 2002US-0373884P.

PR 22-APR-2002; 2002US-0374977P.

PR 16-MAY-2002; 2002US-0381037P.

PR 16-MAY-2002; 2002US-0381039P.

PR 16-MAY-2002; 2002US-0381042P.

PR 17-MAY-2002; 2002US-0381642P.

PR 28-MAY-2002; 2002US-0383656P.

PR 29-MAY-2002; 2002US-0383831P.

PR 25-JUN-2002; 2002US-0391335P.

PR 01-OCT-2002; 2002US-00262511.

XX

PA (CURA-) CURAGEN CORP.

XX

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;

PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;

PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;

PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;

PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;

PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;

XX

XX WPI; 2003-381626/36.

DR N-PSDB; ADA05679.

XX

XX New NOVX polypeptides and nucleic acids, useful for diagnosing,

PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,

PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or

PT pharmacogenomics.

XX

XX Claim 1; Page 135; 586pp; English.

XX

XX The present invention describes NOVX proteins, where X can be 1 to 55

CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide

CC described above and a carrier; (2) a kit comprising, in one or more

CC containers, the composition described above; (3) an isolated nucleic acid

CC molecule which encodes a NOVX protein of the invention; (4) a vector

CC comprising the nucleic acid molecule described above; (5) a cell

CC comprising the above vector; (6) an antibody that immunospecifically

CC binds to the polypeptide described above; (7) methods for determining the

CC presence or amount of the above polypeptide or nucleic acid molecule in a

CC sample; (8) methods for determining the presence of or predisposition to

CC a disease associated with altered levels of expression of the above

CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a

CC method of identifying an agent that binds to the polypeptide described

CC above; (10) a method for identifying a potential therapeutic agent for

CC use in treating a pathology that is related to an aberrant expression or

CC aberrant physiological interactions of the polypeptide; (11) a method of

CC screening for a modulator of activity or of latency or predisposition to

CC a pathology associated with the polypeptide; (12) a method for modulating

CC the activity of the polypeptide described above; (13) methods of treating

CC or preventing a pathology associated with the above polypeptide in a

CC mammal; and (14) a method for producing the above polypeptide. NOVX

CC sequences have antidiabetic, anorectic, antibacterial, virucide,

CC immunomodulator, cytosatic, nootropic, neuroprotective, antiparkinsonian

CC and antilipaeic activities, and can be used in gene therapy. The

CC polypeptide is useful in manufacturing a medicament for treating a

CC syndrome associated with a human disease. The polypeptide or the nucleic

CC acid molecule may be used to diagnose, treat or prevent metabolic

CC disorders such as diabetes or obesity, infections, cachexia, cancer,

CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's

CC disease, immune disorders, haematopoietic disorders and various

CC dyslipidaemias. The nucleic acids can also be used as hybridisation

CC probes, in chromosome mapping, tissue typing, preventive medicine and

CC pharmacogenomics. The present sequence represents a human NOVX from the

CC present invention.

XX

SQ Sequence 537 AA;

Query Match 95.5%; Score 2888; DB 6; Length 537;

Best Local Similarity 100.0%; Pred. No. 1.3e-269;

Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 MGATGAEEPLQSLVWVQORCAVSLPARALLRWRSFPGAGAPGADACVVPVSEIIAV 85

DB 1 MGATGAEEPLQSLVWVQORCAVSLPARALLRWRSFPGAGAPGADACVVPVSEIIAV 60

QY 86 EETDVHGKHGSGKWKQMEKPYAFTVHCVKRARRHWKWAQVTFWCPEQLCHLWQLTLR 145

DB 61 EETDVHGKHGSGKWKQMEKPYAFTVHCVKRARRHWKWAQVTFWCPEQLCHLWQLTLR 120

QY 146 EMLEKLTSRPKHLVFNPFPGKGQGRKIYERKVAFLFTLASITTDIIIVTSHANQAKETL 205

DB 121 EMLEKLTSRPKHLVFNPFPGKGQGRKIYERKVAFLFTLASITTDIIIVTSHANQAKETL 180

QY 206 YEINIDKYDGIVCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRVLPVPSLRIGIIPAGS 265

DB 181 YEINIDKYDGIVCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRVLPVPSLRIGIIPAGS 240

QY 266 TDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGAGYFGYGDIIKDS 325

DB 241 TDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGAGYFGYGDIIKDS 300

QY 326 EKKRWLGLARYDFSGSLKTLFSLHHCYEGTVSFPLPAQHTVGSPPDRKPCPRAGCFVCRQSQK 385

DB 301 EKKRWLGLARYDFSGSLKTLFSLHHCYEGTVSFPLPAQHTVGSPPDRKPCPRAGCFVCRQSQK 360

QY 386 LEEBQKALYGLEAAEDVEEMQVCGKFLATNATNMSCACRRSPRGLSPAHLGSSDL 445

DB 361 LEEBQKALYGLEAAEDVEEMQVCGKFLATNATNMSCACRRSPRGLSPAHLGSSDL 420

QY 446 ILIRKCSRFNPLRLIRHTNQDQDFTFVEVYRVKXQFQTSKHEMEDSDSLKEGKGKRF 505

DB 421 ILIRKCSRFNPLRLIRHTNQDQDFTFVEVYRVKXQFQTSKHEMEDSDSLKEGKGKRF 480

QY 506 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLPARGIEENPKPDSHS 562

DB 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLPARGIEENPKPDSHS 537

RESULT 4



28-MAY-2002; 2002US-0383656P.  
 29-MAY-2002; 2002US-0383831P.  
 25-JUN-2002; 2002US-0391335P.

(SMIT/) SMITHSON G.  
 (MILL/) MILLET I.  
 (PEYM/) PEYMAN J. A.  
 (KEKU/) KEKUDA R.  
 (JUJU/) JU J.  
 (LILL/) LI L.  
 (GUOX/) GUO X.  
 (PATT/) PATTURAJAN M.  
 (SPYT/) SPYTEK K. A.  
 (EDIN/) EDINGER S. R.  
 (ELLE/) ELLERMAN K.  
 (MALY/) MALYANKAR U. M.  
 (ORTT/) ORT T.  
 (GORM/) GORMAN L.  
 (ZERH/) ZERHUSEN B. D.  
 (ANDE/) ANDERSON D. W.  
 (ZHON/) ZHONG M.  
 (CATT/) CATTERTON E.  
 (JIWW/) JI W.  
 (MILL/) MILLER C. E.  
 (RAST/) RASTELLI L.  
 (STON/) STONE D. J.  
 (PENA/) PENA C. E. A.  
 (SHEN/) SHENOY S. G.  
 (SHIM/) SHIMKETS R. A.  
 (ROTH/) ROTHENBERG M. E.  
 (LEAC/) LEACH M. D.  
 (AGEE/) AGEE M. L.  
 (BERG/) BERGHS C.  
 (DIP1/) DIPIPPO V. A.  
 (EISE/) EISEN A.  
 (GANG/) GANGOLLI E. A.  
 (RIEG/) RIEGER D. K.  
 (SPAD/) SPADERNA S. K.

Smithson G., Millet I., Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 Paturajan M, Spyttek KA, Edinger SR, Ellerman K, Malyankar UM;  
 Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
 Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
 Shmkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 Eisen A, Gangolli EA, Rieger DK, Spaderna SK;  
 MPI; 2004-213931/20.  
 N-PSDB; ADN62844.

Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
 diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.

Claim 1; SEQ ID NO 40; 395pp; English.

The invention relates to isolated NOVX polypeptides and polynucleotides.  
 NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
 treat a medical condition in human related to the aberrant expression and  
 activity of NOVX polypeptides. For example, NOVX polypeptides and  
 polynucleotides may be used to treat disorders associated with decreased  
 expression or activity of NOVX by supplementing the patient our  
 production or to rectify mutations. Conversely, antisense NA molecules  
 may be administered to down regulate expression of NOVX polypeptides by  
 binding with the cells own genes and preventing their expression. NOVX  
 polynucleotides and complementary sequences may also be used as DNA  
 probes in diagnostic assays to detect and quantitate the presence of  
 similar sequences in samples, and so which patients may be in need of  
 restorative therapy. NOVX polypeptides may also be used as antigens in  
 the production of antibodies and in assays to identify modulators  
 (agonists and antagonists) of the expression and activity of NOVX. The  
 anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
 used to modulate NOVX polynucleotide expression and activity of NOVX  
 polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
 diagnostic agents for detecting the presence of NOVX in samples. NOVX

polypeptides and polynucleotides may be used in this way to prevent,  
 diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
 disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
 disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,  
 haematopoietic disorders, and the various dyslipidaemias, metabolic  
 disturbances associated with obesity, the metabolic syndrome X and  
 wasting disorders associated with chronic diseases and various cancers.  
 They may also be used as antibacterial agents. The present sequence  
 represents the amino acid sequence of a human NOVX protein.

Sequence 537 AA;

Query Match 95.5%; Score 2888; DB 8; Length 537;  
 Best Local Similarity 100.0%; Prod. No. 1.3e-269;  
 Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 MGATGAAPLQSLVWVKQRCVAVSLPARALLRWRSPGCGAGAPGADACSVPYSEIIV 85  
 Db 1 MGATGAAPLQSLVWVKQRCVAVSLPARALLRWRSPGCGAGAPGADACSVPYSEIIV 60  
 Qy 86 EETDVHGKQSGKWKQKWEKPYAFTVHCVRARHRWKWAQVTFWCPBEQLCHLWLT 145  
 Db 61 EETDVHGKQSGKWKQKWEKPYAFTVHCVRARHRWKWAQVTFWCPBEQLCHLWLT 120  
 Qy 146 EMLEKLTSPRKHLVFINPFGKQGRKRIYERKVAPLFTLASITTDIIVTETHANQAKETL 205  
 Db 121 EMLEKLTSPRKHLVFINPFGKQGRKRIYERKVAPLFTLASITTDIIVTETHANQAKETL 180  
 Qy 206 YEINIDKYGIVCGGDMFSEVLHGLIGRTORSAGVDQNHPRVLPSSURIGIIPAGS 265  
 Db 181 YEINIDKYGIVCGGDMFSEVLHGLIGRTORSAGVDQNHPRVLPSSURIGIIPAGS 240  
 Qy 266 TDCVCYSTVTGSDAETSALHIVGDSLAMDVSSVHNSHTLLRYSVLLGYGYGDIKDS 325  
 Db 241 TDCVCYSTVTGSDAETSALHIVGDSLAMDVSSVHNSHTLLRYSVLLGYGYGDIKDS 300  
 Qy 326 EKKRWGLGARYDFSLGKTLFSLHHCYEGTVSFLPAQHTVGSPPDRKPCRAGFCVCRSQKQ 385  
 Db 301 EKKRWGLGARYDFSLGKTLFSLHHCYEGTVSFLPAQHTVGSPPDRKPCRAGFCVCRSQKQ 360  
 Qy 386 LEEBOKKALVLEAAEDVEWQVCGFLAINATNMSCACRRSPRGLSPAHLGDSDDL 445  
 Db 361 LEEBOKKALVLEAAEDVEWQVCGFLAINATNMSCACRRSPRGLSPAHLGDSDDL 420  
 Qy 446 ILIRKCSRFLRLIRHTNQDQDFTFVEYVRVKKQFTSKHMEDESDLEKGGKKRF 505  
 Db 421 ILIRKCSRFLRLIRHTNQDQDFTFVEYVRVKKQFTSKHMEDESDLEKGGKKRF 480  
 Qy 506 GHICSSHPSCCTVSNSSWNCDEVLHSPAEIVRVHCOLVRLFARGIEENPKPDSHS 562  
 Db 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAEIVRVHCOLVRLFARGIEENPKPDSHS 537

RESULT 6  
 AAM49115  
 ID AAM49115 standard; protein; 537 AA.  
 AC AAM49115;  
 XX  
 DT 20-MAY-2002 (first entry)  
 XX  
 DE Human ceramide kinase hCERK1.  
 XX  
 KW Human; ceramide kinase; hCERK1; drug screening; gene therapy;  
 KW neurological disease; inflammation; human immunodeficiency virus;  
 KW HIV infection; type 2 diabetes; obesity; sepsis; arteriosclerosis;  
 KW cancer; neuroprotective; antiinflammatory; anti-HIV; antidiabetic;  
 KW anorectic; antibacterial; antiseptic; antiarteriosclerotic; cyostatic;  
 KW enzyme.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200196575-A1.

```
XX 20-DEC-2001.
XX 11-JUN-2001; 2001WO-JP004889.
XX 14-JUN-2000; 2000JP-00178039.
XX (SANY ) SANKYO CO LTD.
XX Sugiura M, Kono K, Kohama T;
XX WPI; 2002-179513/23.
XX DR N-PSDB; ABA96945.
XX Human ceramide kinase gene and the enzyme encoded by it for screening
XX substances as drugs for neurological, inflammatory and other disorders.
XX Claim 1; Page 54-57; 61pp; Japanese.
XX This sequence represents a human ceramide kinase designated hCKER1. The
XX invention relates to hCKER1, nucleic acids encoding it, expression
XX vectors and host cells containing hCKER1 nucleic acids, the recombinant
XX production of hCKER1 and antibodies specific for hCKER1. The invention
XX also encompasses methods of isolating hCKER1 from samples, the use of
XX hCKER1 in drug screening, and the use of hCKER1 nucleic acid sequences in
XX gene therapy. hCKER1 mediates the ATP-dependent 1-phosphorylation of
XX ceramides and can be used to screen for therapeutic and preventive agents
XX for a wide range of disorders. Such disorders include neurological
XX disease, inflammation, human immunodeficiency virus (HIV) infection, type
XX 2 diabetes, obesity, sepsis, arteriosclerosis and cancer
XX
XX Sequence 537 AA;
XX
XX Query Match 95.2%; Score 2880; DB 5; Length 537;
XX Best Local Similarity 99.6%; Pred. No. 7.7e-269;
XX Matches 535; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX 26 MGATGAAPQLQSVLWVKQORCAVSLPARALLRWRSFGAGAPGADACSVPVSEIIAV 85
XX 1 MGATGAAPQLQSVLWVKQORCAVSLPARALLRWRSFGAGAPGADACSVPVSEIIAV 60
XX
XX 86 EETDVHGKGQSGKQWQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPEQOLCHLWLT 145
XX 61 EETDVHGKGQSGKQWQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPEQOLCHLWLT 120
XX
XX 146 EMLEKLTSRPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL 205
XX 121 EMLEKLTSRPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL 180
XX
XX 206 YEINIDKYDGI VCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPAGS 265
XX 181 YEINIDKYDGI VCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPAGP 240
XX
XX 266 TDCVCYSTVGTSDAETSALHIVGDSLAMDVSSVHHNSTLLRYSVSLLGFGYGDIIKDS 325
XX 241 TDCVCYSTVGTSDAETSALHIVGDSLAMDVSSVHHNSTLLRYSVSLLGFGYGDIIKDS 300
XX
XX 326 EKKWGLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCRSQKQ 385
XX 301 EKKWGLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCRSQKQ 360
XX
XX 386 LEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 445
XX 361 LEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
XX
XX 446 ILIRKCSRNFRLRFLIRHTNQDQDFTFVEVYRVKQFQFTSKHMEDESDSLKGGKKRF 505
XX 421 ILIRKCSRNFRLRFLIRHTNQDQDFTFVEVYRVKQFQFTSKHMEDESDSLKGGKKRF 480
XX
XX 506 GHICSSHPSCCCTVSNSSWNCDEVLHSPATEVHVHCOLVELFARGLEENPKPDSHS 562
XX 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPATEVHVHCOLVELFARGLEENPKPDSHS 537
```

---

```
RESULT 7
ABR56302
ID ABR56302 standard; protein; 481 AA.
XX
AC ABR56302;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human Sphingosine kinase 4-related protein.
XX
KW Human; enzyme; haemostatic; sphingosine kinase 4; SPHK4;
XX platelet transfusion; platelet stabiliser.
XX
OS Homo sapiens.
XX
PN WO2003031627-A1.
XX
PD 17-APR-2003.
XX
PF 28-SEP-2001; 2001WO-JP008537.
XX
PR 28-SEP-2001; 2001WO-JP008537.
XX (HOKK-) HOKKAIDO TECHNOLOGY LICENSING OFFICE CO.
XX (CHBI-) CHEM BIOLOGY INST.
XX
PI Igaraashi Y, Kihara A;
XX
DR WPI; 2003-354917/33.
XX
XX Platelet derived polypeptides with sphingosine kinase activity for
XX treatment of sphingosine related disorders.
XX
XX Example 4; Fig 4; 39pp; Japanese.
XX
XX The present invention relates to human sphingosine kinase 4 (SPHK4;
XX ABR56301). The kinase can be used for the diagnosis and treatment of
XX sphingosine related disorders. The kinase can also be potentially used
XX for controlling toxicity of platelet transfusion and as a platelet
XX stabiliser. The present sequence was used to illustrate the invention
XX
XX Sequence 481 AA;
XX
XX Query Match 85.6%; Score 2588; DB 6; Length 481;
XX Best Local Similarity 100.0%; Pred. No. 1.1e-240;
XX Matches 481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 82 IIAVEETDVHGKGQSGKQWQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPEQOLCHLWL 141
XX 1 IIAVEETDVHGKGQSGKQWQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPEQOLCHLWL 60
XX
XX 142 QTLREMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQA 201
XX 61 QTLREMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQA 120
XX
XX 202 KETLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGI 261
XX 121 KETLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGI 180
XX
XX 262 PAGSTDCVCYSTVGTSDAETSALHIVVGDSDLAMDVSSVHHNSTLLRYSVSLLGFGYGD 321
XX 181 PAGSTDCVCYSTVGTSDAETSALHIVVGDSDLAMDVSSVHHNSTLLRYSVSLLGFGYGD 240
XX
XX 322 IKDSEKKRWGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCRQ 381
XX 241 IKDSEKKRWGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCRQ 300
XX
XX 382 SKQOLEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDG 441
XX 301 SKQOLEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDG 360
XX
XX 442 SSDLILIRKCSRNFRLRFLIRHTNQDQDFTFVEVYRVKQFQFTSKHMEDESDSLKEGG 501
```

Db 361 SDDLILIRKCSRFNLFIRHTNQDQDFTFVEVYRVKQFTSKMEDESDLKEGG 420  
 QY 502 KKRFGHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVFLFARGIEENPKPDH 561  
 Db 421 KKRFGHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVFLFARGIEENPKPDH 480  
 QY 562 S 562  
 Db 481 S 481

## RESULT 8

AAV96059

ID AAY96059 standard; protein; 460 AA.

XX AC

XX AAY96059;

XX 05-DEC-2000 (first entry)

XX Human sphingosine kinase C.

XX Sphingosine kinase C; SKC; human; drug screening; infection;

KW antiinflammatory; antiallergic; anticancer; inflammation; allergy;

KW cancer; therapy; diagnosis.

XX Homo sapiens.

OS WO200052173-A2.

XX 08-SEP-2000.

XX 02-MAR-2000; 2000WO-CA000223.

XX 02-MAR-1999; 99US-0122516P.

XX (ALIX ) NPS ALLELIX CORP.

XX Munroe D, Gupta A, Falzone GR;

PI WPI; 2000-572185/53.

DR N-PSDB; AAA50510.

XX New human sphingosine kinase A, B and C polynucleotides and polypeptides useful in e.g. chromosome and gene mapping, and detecting inflammation or disease associated with abnormal levels of sphingosine kinase expression.

XX Disclosure; Fig 9; 81pp; English.

XX The present sequence is that of human sphingosine kinase C (SKC), an enzyme that phosphorylates sphingosine to form sphingosine 1-phosphate. The sequence was deduced from that of a polynucleotide (see AAA50510) isolated from an HeLa cDNA library. The invention provides polynucleotides (see AAA50508-10) and polypeptides (see AAY96057-59) for the human sphingosine kinase (SK) homologues SKA, SKB and SKC. The polypeptides can be obtained using recombinant DNA methods, and host cells containing expression vectors including SK polynucleotides are used in a claimed method of screening for compounds that inhibit or activate human SK activity. Human SK specific antibodies, inhibitors, ligands or their analogues can be used as bioactive agents to treat inflammation or disease including viral, bacterial or fungal infections, allergic responses, mechanical injury associated with trauma, hereditary diseases, lymphoma or carcinoma, and other conditions with activate the genes of kidney, lung, heart, lymphoid or tissues of the nervous system

SQ Sequence 460 AA;

Query Match 81.4%; Score 2463; DB 3; Length 460;

Best Local Similarity 99.6%; Pred. NO. 1.2e-228;

Matches 458; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 103 MEKPYAFVHCVRARRHWKWAQVFWCPPEQLCHLWQLTRMLEKLTSRPKHLVFI 162  
 |||||

Db 1 MEKPYAFVHCVRARRHWKWAQVFWCPPEQLCHLWQLTRMLEKLTSRPKHLVFI 60  
 QY 163 NPFGGKGGKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLYEINIDKYDGI VCVGGD 222  
 |||||  
 Db 61 NPFGGKGGKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLYEINIDKYDGI VCVGGD 120  
 |||||  
 QY 223 GMFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPAGSTDCVCYSTVGTSDAETS 282  
 |||||  
 Db 121 GMFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPAGSTDCVCYSTVGTSDAETS 180  
 |||||  
 QY 283 ALHIVVGDLSAMDVSSVHHNSTLLRYSVSLIGYFYGDIIKDSKKRWGLARVDFSLK 342  
 |||||  
 Db 181 ALHIVVGDLSAMDVSSVHHNSTLLRYSVSLIGYFYGDIIKDSKKRWGLARVDFSLK 240  
 |||||  
 QY 343 TFLSHHCYEGTVSFLPAQHTVGSRRKPCRAGFCVCRQSKQLEEEOKKALYLEAAED 402  
 |||||  
 Db 241 TFLSHHCYEGTVSFLPAQHTVGSRRKPCRAGFCVCRQSKQLEEEOKKALYLEAAED 300  
 |||||  
 QY 403 VEEQVVCVKFLAINATNMSCACRRSPRGLSPAHLGDSDDLILIRKCSRFNLFILIR 462  
 |||||  
 Db 301 VEEQVVCVKFLAINATNMSCACRRSPRGLSPAHLGDSDDLILIRKCSRFNLFILIR 360  
 |||||  
 QY 463 HTNQDQDFTFVEVYRVKQFTSKMEDESDLKEGGKRFHICSSHPSCCCTVSN 522  
 |||||  
 Db 361 HTNQDQDFTFVEVYRVKQFTSKMEDESDLKEGGKRFHICSSHPSCCCTVSN 420  
 |||||  
 QY 523 SWNCDEVLHSPAIEVRVHCOLVFLFARGIEENPKPDHSHS 562  
 |||||  
 Db 421 SWNCDEVLHSPAIEVRVHCOLVFLFARGIEENPKPDHSHS 460  
 |||||

## RESULT 9

AAE07884

ID AAE07884 standard; protein; 471 AA.

XX AC

XX AAE07884;

XX 01-NOV-2001 (first entry)

XX Human sphingosine kinase (SphK) protein #2.

XX Human; sphingosine kinase; SphK; restenosis; ischaemia; gene therapy;

KW antisense therapy; cancer; sphingolipid; signalling molecule; apoptosis;

KW cytosolic; tumour necrosis factor-alpha; TNF; atherosclerosis; lymphoma;

KW leukaemia; vasotropic; cell proliferative disorder; vascular disease.

XX Homo sapiens.

OS WO200160990-A2.

XX 23-AUG-2001.

XX 14-FEB-2001; 2001WO-US004789.

XX 14-FEB-2000; 2000US-0182360P.

XX 22-MAR-2000; 2000US-0191261P.

XX (CURA-) CURAGEN CORP.

XX (GETH ) GENENTECH INC.

XX Rastelli L;

PI WPI; 2001-514770/56.

DR N-PSDB; AAD14426.

XX An isolated sphingosine kinase polypeptide useful for treating a SphK-

XX associated disorder especially cancer, restenosis or ischemia in a human.

XX Claim 1; Fig 1; 107pp; English.

XX The present invention relates to sphingosine kinase (SphK) polypeptides

XX and nucleic acids encoding them. SphK is useful for treating a SphK-

XX associated disorder especially cancers such as leukaemia, lymphoma,

CC ovarian, breast, lung, colon, testicular, stomach and skin,  
CC atherosclerosis, restenosis or ischaemia and cell proliferative disease  
CC or disorder associated with vascular diseases. Spk gene is used in gene  
CC therapy and antisense-therapy. Sphingolipids serving as signalling  
CC molecules, have recently emerged as regulators of cell growth,  
CC differentiation, diverse cell phenotypes and cell death. Activation of  
CC Spk by tumour necrosis factor (TNF)-alpha inhibits apoptosis in human  
CC endothelial cells. The present sequence is human sphingosine kinase  
CC (Spk) protein  
XX  
SQ Sequence 471 AA;

Query Match 81.2%; Score 2456.5; DB 4; Length 471;  
Best Local Similarity 97.5%; Pred. No. 5.4e-228;  
Matches 459; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

Qy 103 MEKPYATVHCVGKARRHRWKQAQVTFWCPPEQCHLWQLTREMLEKLSRPHKLLVFI 162  
Db 1 MEKPYATVHCVGKARRHRWKQAQVTFWCPPEQCHLWQLTREMLEKLSRPHKLLVFI 60

Qy 163 NPFQGGKQKRIYERKVAFLFTLASITTDII-----VTEHANQAKETLYEINID 211  
Db 61 NPFQGGKQKRIYERKVAFLFTLASITTDIIIGNKFYVYVEVITEHANQAKETLYEINID 120

Qy 212 KYDGI VCGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLPSSLRIGIIPAGSTDCVCY 271  
Db 121 KYDGI VCGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLPSSLRIGIIPAGSTDCVCY 180

Qy 272 STVGTSDAETSALHIVVGDLSLMDVSVVHNSLTLLRYSVLLGYGFGYDIIKDEKKRWL 331  
Db 181 STVGTSDAETSALHIVVGDLSLMDVSVVHNSLTLLRYSVLLGYGFGYDIIKDEKKRWL 240

Qy 332 GLARYDFSLKTFLSHCYEGTVSFLPAQHTVSGPRDRKPCRACFCVCRSKQOLEBEQK 391  
Db 241 GLARYDFSLKTFLSHCYEGTVSFLPAQHTVSGPRDRKPCRACFCVCRSKQOLEBEQK 300

Qy 392 KALYGLEAAEDVEEWVQVCGKFLAINATNMSCACRRSPRGLSPAHLHGDGSSDLILIRKC 451  
Db 301 KALYGLEAAEDVEEWVQVCGKFLAINATNMSCACRRSPRGLSPAHLHGDGSSDLILIRKC 360

Qy 452 SRNFLRLRHTNQDQDFTFVEVYRVKFKQTSKHMEDESDLEGGKKRFGHTCSS 511  
Db 361 SRNFLRLRHTNQDQDFTFVEVYRVKFKQTSKHMEDESDLEGGKKRFGHTCSS 420

Qy 512 HPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKDSDHS 562  
Db 421 HPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKDSDHS 471

RESULT 10  
ADP55248  
ID ADP55248 standard; protein; 531 AA.

XX AC ADP55248;

XX DT 18-NOV-2004 (first entry)

XX DE Human PRO protein sequence SEQ ID NO:1224.

XX KW human; PRO; immune related disease; inflammatory immune response;  
KW immune response stimulation; antiallergic; antianaemic; antiarthritic;  
KW antiasthmatic; antidiabetic; antinflammatory; antipsoriatic;  
KW antirheumatic; antichyroid; CNS; dermatological; gastrointestinal;  
KW haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;  
KW nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;  
KW virucide; gene therapy.

XX OS Homo sapiens.

XX PN WO2004039956-A2.

XX PD 13-MAY-2004.

XX ,

PF 28-OCT-2003; 2003WO-US034381.  
XX  
PR 29-OCT-2002; 2002US-0422472P.  
XX  
XX (GETH ) GENENTECH INC.  
XX Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;  
PI Wood WI, Wu TD;  
XX  
XX WPI: 2004-376182/35.  
DR N-PSDB; ADP55247.  
XX  
XX New PRO polynucleotides and polypeptides, useful in diagnosing  
PT and treating an immune related disease, e.g. systemic lupus  
PT erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in  
PT stimulating an immune response.  
XX  
XX Claim 1; SEQ ID NO 1224; 3009pp; English.  
XX  
XX The present invention describes an isolated PRO nucleic acid (1). Also  
CC described: (1) a vector comprising (1); (2) a host cell comprising the  
CC vector of (1); (3) a process for producing a PRO polypeptide; (4) an  
CC isolated PRO polypeptide; (5) a chimeric molecule comprising the  
CC polypeptide of (4) fused to a heterologous amino acid sequence; (6) an  
CC antibody which specifically binds to a polypeptide of (4); (7) a  
CC composition of matter comprising a polypeptide of (4), an agonist or  
CC antagonist of the polypeptide or an antibody that binds to the  
CC polypeptide in combination with a carrier; (8) an article of manufacture  
CC comprising a container, a label on the container and a composition of  
CC matter of (7); (9) a method of treating an immune related disease in a  
CC mammal; (10) a method for determining the presence of a PRO polypeptide  
CC in a sample suspected of having the polypeptide; (11) a method of  
CC diagnosing an immune related disease or an inflammatory immune response  
CC in a mammal; (12) a method of identifying a compound that inhibits or  
CC mimics the activity of or expression of a gene encoding a PRO polypeptide  
CC ; and (13) a method of stimulating the immune response in a mammal. The  
CC PRO sequences have antiallergic, antianaemic, antiarthritic,  
CC antiasthmatic, antidiabetic, antinflammatory, antipsoriatic,  
CC antirheumatic, antichyroid, CNS, dermatological, gastrointestinal,  
CC haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,  
CC nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and  
CC virucide activities, and can be used in gene therapy. The nucleic acid  
CC (1) and the encoded polypeptides, compositions, kits and methods are  
CC useful in diagnosing and treating an immune related disease and in  
CC stimulating an immune response. The present sequence represents a human  
CC PRO protein from the present invention.  
XX  
SQ Sequence 531 AA;

Query Match 80.3%; Score 2428; DB 8; Length 531;  
Best Local Similarity 85.1%; Pred. No. 3.7e-225;  
Matches 450; Conservative 33; Mismatches 46; Indels 0; Gaps 0;

Qy 26 MGATGAAPLOSIVLVKQRCVSLPAPARALLWRSFGPAGAPGADACVSPVSEIIAV 85  
Db 1 MGAGAAEPLHSLVWVKRRRCVSLPAPARALLWRSFGPAGAPGADACVSPVSEIIAV 60

Qy 86 EETDVHGKQSGKQWQKPYAFTVHCVKRARRHRWKQAQVTFWCPPEQCHLWQLTLR 145  
Db 61 EEKDDCEKHSSGRWHKQENPFATVHRVCRVHRHRWKARVTFWSADEQCHLWQLTLR 120

Qy 146 EMLEKLTSRPKHLLVFINPFGGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL 205  
Db 121 GLLESLSRPKHLLVINPFGGQKRIYERKVAFLFTLASITTEIIITEHANQAKETL 180

Qy 206 YEINIDKYDGI VCGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLPSSLRIGIIPAGS 265  
Db 181 YEINTSDYDGI VCGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLPSSLRIGIIPAGS 240

Qy 266 TDCVCTVGTSDAETSALHIVVGDLSLMDVSVVHNSLTLLRYSVLLGYGFGYDIIKDS 325  
Db 241 TDCVCTVGTNDAETSALHIIIGDSLAIQDVSVVHNTLLRYSVLLGYGFGYDIIKDS 300

```
QY 326 EKRWGLGLARYDFSLGKTLFLSHCYEGTVSFLPAQHTVGSPRDRKPCBAGCFVCRQSKQ 385
DB 301 EKRWGLGLARYDFSLGKTLFLSHCYEGTVSFLPAQHTVGSPRDRKPCBAGCFVCRQSKQ 360
QY 386 LEEEOKALYGLAEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGSGSD 445
DB 361 LEEEOKALYGLAEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGSGSD 420
QY 446 ILIRKCSRFNLFRLIRHTNQODQDFTFVEYVRVKKFQFTSKMEDESDSLKGGKKRP 505
DB 421 ILIRKCSRFNLFRLIRHTNQODQDFTFVEYVRVKKFQFTSKMEDESDSLKGGKKRP 480
QY 506 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEE 554
DB 481 GKICKDRPSCTSASRSWNCDEVLHSPAIEVRVHCOLVRLFARGIEE 529

RESULT 11
ADSL11054
ID ADSL11054 standard; protein; 536 AA.
XX AC ADSL11054;
XX DT 16-DEC-2004 (first entry)
XX DE Human therapeutic protein - SEQ ID 1291.
XX KW antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnery;
XX KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;
XX KW aplastic anaemia; cancer; wound healing; gene therapy.
XX OS Homo sapiens.
XX PN WO2004080148-A2.
XX PD 23-SEP-2004.
XX PF 30-SEP-2003; 2003WO-US030720.
XX PR 02-OCT-2002; 2002US-0416186P.
XX PA (NUVE-) NUVELO INC.
XX FI Tang YT, Asundi V, Ren F, Zhang J, Zhang J, Wehrman T, Wang Z, Ma Y;
XX FI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;
XX DR WPI; 2004-668857/65.
XX DR N-PSDB; ADS10370.
XX PT New polynucleotide, useful in preparing a composition for diagnosing or
XX PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
XX PT aplastic anemia or cancer for promoting wound healing.
XX PS Claim 20; SEQ ID NO 1291; 718pp; English.
XX CC The invention relates to a novel isolated polynucleotide and the encoded
XX CC polypeptide. The molecules of the invention demonstrate antiinflammatory,
XX CC neuroprotective, antianaemic, cytostatic and vulnery activities and may
XX CC be useful in preparing a composition for diagnosing or treating
XX CC inflammatory haematopoietic, immune, neurodegenerative or stem cell
XX CC disorders, such as aplastic anaemia or cancer, as well as for promoting
XX CC wound healing. The molecules may also be utilised during gene therapy
XX CC procedures. The current sequence is that of a human therapeutic protein
XX CC of the invention. The current sequence is not shown explicitly within the
XX CC specification but can be accessed from the WIPO web-site.
XX SQ Sequence 536 AA;

Query Match 77.7%; Score 2350; DB 8; Length 536;
Best Local Similarity 78.1%; Pred. No. 1.3e-217;
Matches 452; Conservative 2; Mismatches 9; Indels 116; Gaps 4;

QY 57 LRWR-----SPGPGAGAGADACSVPVSEIIAETDVHGKHQGGSGKQWMEKPYAFT 110
```

```
DB 1 MEWGRVFCVEMPAGRGV-FTDACSVPVSEIIAETDVHGKHQGGSGKQWMEKPYAFT 59
QY 111 VHCVRARRHRWKAQVTFMCPPEQLCHLMLQTLREMLEKLTSPRQHLVFINPFGGKG 170
DB 60 VHCVRARRHRWKAQVTFMCPPEQLCHLMLQTLREMLEKLTSPRQHLVFINPFGGKG 100
QY 171 GKRIYERKAPLFTLASITTDIIVTEHANOAKETLYEINDKYDGIYVCGDGMFSEVLH 230
DB 101 -----ITEHANOAKETLYEINDKYDGIYVCGDGMFSEVLH 137
QY 231 GLIGRTORSAGVDQNHPRAVLPSSLRIGIIPA----- 263
DB 138 GLIGRTORSAGVDQNHPRAVLPSSLRIGIIPA----- 197
QY 264 -----GSTDCVCYVTVGTSDAETSA 283
DB 198 HTFTLVATLGCERHSHPHFMDERRTGCVHVGPGENAGGLDGDSTDCVCYVTVGTSDAETSA 257
QY 284 LHIIVGDSLAMDVSSVHNSTLLRYSVLLQYGYGDIIDKSEKKRWGLARYDFSLGKT 343
DB 258 LHIIVGDSLAMDVSSVHNSTLLRYSVLLQYGYGDIIDKSEKKRWGLARYDFSLGKT 317
QY 344 FLSHHCYEGTVSFLPAQHTVGSPRDRKPCBAGCFVCRQSKQLEEOKALYGLAEADV 403
DB 318 FLSHHCYEGTVSFLPAQHTVGSPRDRKPCBAGCFVCRQSKQLEEOKALYGLAEADV 377
QY 404 EEMOVVCGKFLAINATNMSCACRRSPRGLSPAHLGSGSDLIILIRKCSRFNLFRLIRH 463
DB 378 EEMOVVCGKFLAINATNMSCACRRSPRGLSPAHLGSGSDLIILIRKCSRFNLFRLIRH 437
QY 464 TNQODQDFTFVEYVRVKKFQFTSKMEDESDSLKGGKKRFGHICSSHPSCCTVSNSS 523
DB 438 TNQODQDFTFVEYVRVKKFQFTSKMEDESDSLKGGKKRFGHICSSHPSCCTVSNSS 497
QY 524 WNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 562
DB 498 WNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 536

RESULT 12
ABG13541
ID ABG13541 standard; protein; 746 AA.
XX AC ABG13541;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #13532.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX FI Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX DR N-PSDB; AAS77728.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
```



responsible for genetic disorders or other traits and to assess biodiversity.

Claim 20; SEQ ID NO 43900; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 746 AA;

Query Match 73.2%; Score 2215; DB 4; Length 746;  
Best Local Similarity 68.7%; Pred. No. 2.3e-204;  
Matches 444; Conservative 12; Mismatches 44; Indels 146; Gaps 9;

Qy 6 GPAPLGVRAPAWRTSPAENGATGAAPLOSVLWVKQOR---CAVSLPARAL----- 56  
Db 124 GPAPTL---PHSSDVPPEARAGDPPSQLLTLHAGLSQRACCGSLVSPVHCLFLVSCP 180  
Qy 57 -----LRWWR-----SPGPGAGAGDACSVPSEIIAIEETDVH 91  
Db 181 ELSVIAGREADAGFRPDSMEWGRVFCVEMPAGRGV-FTDACSVPVSEIIAIEETDVH 239  
Qy 92 GKHGSGKQWQKMEKPYAFTVHCVRARRHRKWAQVTFWCPEQLCHLWLQTLREMLEKL 151  
Db 240 GKHGSGKQWQKMEKPYAFTVHCVRARRHRKWAQVTFWCPEQLCHLWLQTLREMLEKL 299  
Qy 152 TSPKHLVLFNPGGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLYEINID 211  
Db 300 -----ITEHANOAKETLYEINID 317  
Qy 212 KYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLPSSLRIGIIPA----- 263  
Db 318 KYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLPSSLRIGIIPAETHVGPGEN 377  
Qy 264 -----GSTDCVCYSTVGTSDAETSALHIV----- 288  
Db 378 AGGLDGTDCVCYSTVGTSDAETSALHIVGCCPEARKPPASRHTACSGHGQLCLGQOL 437  
Qy 289 -----GDSLAMDVSSVHNSLRLRYSVSLGCGYFGDIKDSEKKRW 330  
Db 438 WNACLEASRLQSRMQSPGDSLAMDVSSVHNSLRLRYSVSLGCGYFGDIKDSEKKRW 497  
Qy 331 LGLARYDFSGUKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQOLBEEQ 390  
Db 498 LGLARYDFSGIKTFLSHHCYEGTVSFLPAQHTVGSPPNRKPCRCAGCFVCRQSKQOLBEEQ 557  
Qy 391 KKALYGLEAAEDVEEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDSGLILIRK 450  
Db 558 KKALYGLEAAEDVEEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDSGLILIRK 617  
Qy 451 CSRNFLRLIRHTNQODQDFTFVEYVRVKKFOFTSKHMEDESDLKEGKKRFGHICS 510  
Db 618 CSRNFLRLIRHTNQODQDFTFVEYVRVKKFLFTSKHMEDESDLKEGKKRFGHICS 677

Qy 511 SHPSCCTVSNSSWNCDEVLHSPHSPAEIVRVHQCQLVRLFARGIENP 556  
Db 678 SHPSCCTVSNSSWNCDEVLHSPHSPAEIVR-----AVVMARGCRKVP 718

RESULT 13  
ABRS6301

ID ABR56301 standard; protein; 416 AA.

XX ABR56301;

XX 20-NOV-2003 (first entry)

DE Human Sphingosine kinase 4.

KW Human; enzyme; haemostatic; sphingosine kinase 4; SPHK4;

KW platelet transfusion; platelet stabiliser.

OS Homo sapiens.

PN WO2003031627-A1.

XX 17-APR-2003.

XX 28-SEP-2001; 2001WO-JP008537.

XX 28-SEP-2001; 2001WO-JP008537.

XX (HOKK-) HOKKAIDO TECHNOLOGY LICENSING OFFICE CO.

PA (CHBI-) CHEM BIOLOGY INST.

PI Igarashi Y, Kihara A;

DR WPI; 2003-354917/33.

DR N-PSDB; ACC70838.

PT Platelet derived polypeptides with sphingosine kinase activity for treatment of sphingosine related disorders.

PS Claim 2; Page 29-30; 39pp; Japanese.

CC The present sequence is the protein sequence for human sphingosine kinase 4 (SPHK4). The kinase can be used for the diagnosis and treatment of sphingosine related disorders. The kinase can also be potentially used for controlling toxicity of platelet transfusion and as a platelet stabiliser

XX Sequence 416 AA;

Query Match 73.1%; Score 2210; DB 6; Length 416;  
Best Local Similarity 100.0%; Pred. No. 3e-204;  
Matches 416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 147 MLEKLSRPRKHLVFNPFPGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLY 206  
Db 1 MLEKLSRPRKHLVFNPFPGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLY 60

Qy 207 EINDIKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLPSSLRIGIIPAGST 266

Db 61 EINDIKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLPSSLRIGIIPAGST 120

Qy 267 DCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHNSLRLRYSVSLGCGYFGDIKDSE 326

Db 121 DCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHNSLRLRYSVSLGCGYFGDIKDSE 180

Qy 327 KRWLGLARYDFSGIKTFLSHHCYEGTVSFLPAQHTVGSPPNRKPCRCAGCFVCRQSKQOL 386

Db 181 KRWLGLARYDFSGIKTFLSHHCYEGTVSFLPAQHTVGSPPNRKPCRCAGCFVCRQSKQOL 240

Qy 387 EEOKKALYGLEAAEDVEEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDSGLI 446

Db 241 EEOKKALYGLEAAEDVEEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDSGLI 300

QY 447 LIRKSRNRLFLIRHTNQDQDFTFVYVRVKKFOFTSKHMEDESDLKEGGKKEFG 506  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 301 LIRKSRNRLFLIRHTNQDQDFTFVYVRVKKFOFTSKHMEDESDLKEGGKKEFG 360  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 507 HICSHHPSCCTVSNSNCDGVLHSPAEIVRVHQCQLVRLFARGIENPKPDGSHS 562  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 361 HICSHHPSCCTVSNSNCDGVLHSPAEIVRVHQCQLVRLFARGIENPKPDGSHS 416  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

## RESULT 14

ABGI3543  
 ID ABGI3543 standard; protein; 727 AA.

XX AC ABGI3543;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #13534.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS77730.

XX PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.

XX PS Claim 20; SEQ ID NO 43902; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 727 AA;

Query Match 56.7%; Score 1714.5; DB 4; Length 727;  
 Best Local Similarity 63.7%; Pred. No. 5.2e-156;

Matches	345;	Conservative	0;	Mismatches	0;	Indels	197;	Gaps	3;
QY	195	TEHANOAKETLYEINIDKYDG-							215
Db	1	TEHANOAKETLYEINIDKYDCYTAHSRMPNHNHSPAGKAAPVQGVGYGEGRAFWPWP							60
QY	216	-----IVCVCGDGMFSEVLHGL							232
Db	61	GPQMPAKGOVCRLPLKLEASGLLRSEGRCTCRSPDLRCSSCSIVCVGDDGMFSEVLHGL							120
QY	233	IGRTORSAGVDQNHPRAVLVPSSLRIGIIPA-							263
Db	121	IGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGFANDLAGDEVSVLSGPVSGPEGGMVHT							180
QY	264	-----GSTDCVCYSTVGTSDAETSALH							285
Db	181	FTLVLTALGCEHRSHPHFMDERRTGEHVHVGPNAGGLDGGTDCVCYSTVGTSDAETSALH							240
QY	286	IIV-----GDSLAMDVS							297
Db	241	IIVGCCPEARKPPASRHTACSGHGQLCLGQLWNACLCEASRLQSRMQSPGDSLAMDVS							300
QY	298	SVHNSLTLLRYSVSLLGYGFYGDIIKDSEKKRWLGLARYDFSLKTLFLSHHCYEGTVSFL							357
Db	301	SVHNSLTLLRYSVSLLGYGFYGDIIKDSEKKRWLGLARYDFSLKTLFLSHHCYEGTVSFL							360
QY	358	PAQTVGSPDRKPCRAGCFVCRQSKQLEBEQKALYGLAAEDVEWVVCCKFLAIN							417
Db	361	PAQTVGSPDRKPCRAGCFVCRQSKQLEBEQKALYGLAAEDVEWVVCCKFLAIN							420
QY	418	ATNMSCACRRSPRGLSPAHLGDGSSDILIRKCSRNFRLRFLIRHTNQDQDFTFVEV							477
Db	421	ATNMSCACRRSPRGLSPAHLGDGSSDILIRKCSRNFRLRFLIRHTNQDQDFTFVEV							480
QY	478	YRVKKFQFTSKHMEDESDLKEGGKKEFGHICSSHPSCCTVSNSNCDGVLHSPAE							537
Db	481	YRVKKFQFTSKHMEDESDLKEGGKKEFGHICSSHPSCCTVSNSNCDGVLHSPAE							540
QY	538	VR 539							
Db	541	VR 542							
RESULT 15									
ABB07854	ID	ABB07854	standard; protein; 326 AA.						
XX AC	ABB07854;								
XX DT	03-JUL-2002 (first entry)								
XX DE	Human sphingosine kinase-like protein.								
XX KW	Human sphingosine kinase-like protein; intracellular signalling; cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma; autoimmune disease; rheumatoid arthritis; Parkinson's disease.								
XX OS	Homo sapiens.								
XX PN	WO200228906-A2.								
XX PD	11-APR-2002.								
XX PF	05-OCT-2001; 2001WO-EP011516.								
XX PR	06-OCT-2000; 2000US-0238005P.								
XX PR	23-AUG-2001; 2001US-0314113P.								
XX PA	(FARB ) BAYER AG.								
XX PI	Kossida S, Encinas J;								
XX DR	WPI; 2002-340094/37.								

DR N-PSDB; ABL40822.  
XX New reagent for modulating the activity of sphingosine kinase-like  
PT protein polypeptide or polynucleotide and treating cancer, asthma,  
PT allergy, an autoimmune disease, or a central or peripheral nervous system  
XX disorder.  
XX  
XX  
PS Claim 25; Fig 2; 120pp; English.  
XX  
XX  
CC The invention relates to a human sphingosine kinase-like protein. The  
CC polypeptide can be expressed by standard recombinant methodology. The  
CC sphingosine kinase-like protein and gene can be used to regulate  
CC intracellular signalling and consequently cell proliferation and  
CC apoptosis. Such regulation is useful for treating cancer, allergies (e.g.  
CC asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and  
CC peripheral nervous system disorders (e.g. Parkinson's disease). The  
CC present sequence represents the human sphingosine kinase-like protein  
XX  
SQ Sequence 326 AA;  
Query Match 54.2%; Score 1640.5; DB 5; Length 326;  
Best Local Similarity 96.3%; Pred. No. 2.3e-149;  
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;  
Qy 155 PKHLLVFINPRGKGQKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 203  
Db 1 PKHLLVFINPRGKGQKRIYERKVAFLFTLASITTDIIIGNKFVYVYVEVITEHANOAKE 60  
Qy 204 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPA 263  
Db 61 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPA 120  
Qy 264 GSTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSVVHNSLTLRYSVSLGCGFYGDIIK 323  
Db 121 GSTDCVCYSTVGTSDAETSALHIVVGDLSLMDVSVVHNSLTLRYSVSLGCGFYGDIIK 180  
Qy 324 DSEKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCRQSK 383  
Db 181 DSEKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCRQSK 240  
Qy 384 QQLBEEQKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 443  
Db 241 QQLBEEQKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300  
Qy 444 DLILIRKCSRFNLFRLIRHTNQDDQ 469  
Db 301 DLILIRKCSRFNLFRLIRHTNQDDQ 326

Search completed: September 3, 2005, 04:37:42  
Job time : 116.766 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 04:08:02 ; Search time 111.217 Seconds  
(without alignments)  
1990.064 Million cell updates/sec

Title: US-10-631-958-11

Perfect score: 3025

Sequence: 1 HEANGAPAGLGVAPPAAWRT.....QLVRLFARGIEENKPDHS 562

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1774312 seqs, 393823214 residues

Total number of hits satisfying chosen parameters: 1774312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/FCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3025	100.0	562	10	US-09-969-896-11
2	3025	100.0	562	16	US-10-631-958-11
3	2888	95.5	537	10	US-09-969-896-10
4	2888	95.5	537	15	US-10-262-511-40
5	2888	95.5	537	16	US-10-631-958-10
6	2888	95.5	537	16	US-10-618-941-121
7	2880	95.2	537	14	US-10-315-597A-2
8	2456.5	81.2	471	9	US-09-784-810A-6
9	2456.5	81.2	471	18	US-10-876-281-6
10	1640.5	54.2	326	9	US-09-784-810A-11
11	1640.5	54.2	326	10	US-09-969-896-2

12	1640.5	54.2	326	16	US-10-631-958-2	Sequence 2, Appli
13	1640.5	54.2	326	18	US-10-876-281-11	Sequence 11, Appl
14	986	32.6	182	9	US-09-784-810A-29	Sequence 29, Appl
15	986	32.6	182	18	US-10-876-281-29	Sequence 29, Appl
16	647	21.4	144	9	US-09-784-810A-8	Sequence 8, Appli
17	647	21.4	144	18	US-10-876-281-8	Sequence 8, Appli
18	611.5	20.8	588	16	US-10-425-115-361076	Sequence 361076,
19	597.5	19.8	575	16	US-10-425-115-270874	Sequence 270874,
20	480.5	15.9	613	16	US-10-437-963-195284	Sequence 195284,
21	404.5	13.4	484	16	US-10-437-963-133133	Sequence 133133,
22	389	12.9	487	15	US-10-424-599-190238	Sequence 190238,
23	362.5	12.0	760	16	US-10-425-115-213665	Sequence 213665,
24	357.5	11.8	490	16	US-10-425-115-249553	Sequence 249553,
25	344.5	11.4	381	16	US-10-425-115-249555	Sequence 249555,
26	341	11.3	520	15	US-10-424-599-202261	Sequence 202261,
27	335	11.1	79	9	US-09-784-810A-10	Sequence 10, Appl
28	335	11.1	79	18	US-10-876-281-10	Sequence 10, Appl
29	334	11.0	245	15	US-10-425-114-53957	Sequence 53957, A
30	331.5	11.0	446	16	US-10-437-963-197631	Sequence 197631,
31	330.5	10.9	757	16	US-10-437-963-180379	Sequence 180379,
32	326.5	10.8	470	16	US-10-425-115-282512	Sequence 282512,
33	326.5	10.8	521	15	US-10-425-114-68423	Sequence 68423, A
34	313	10.3	552	16	US-10-437-963-161184	Sequence 161184,
35	313	10.3	677	16	US-10-425-115-264087	Sequence 264087,
36	312.5	10.3	373	16	US-10-767-701-44927	Sequence 44927, A
37	310.5	10.3	641	15	US-10-348-052-28	Sequence 28, Appl
38	310.5	10.3	641	16	US-10-622-011-28	Sequence 28, Appl
39	306	10.1	490	14	US-10-053-510-19	Sequence 19, Appl
40	306	10.1	490	15	US-10-348-052-19	Sequence 19, Appl
41	306	10.1	490	16	US-10-622-011-19	Sequence 19, Appl
42	301	10.0	524	14	US-10-053-510-20	Sequence 20, Appl
43	301	10.0	524	15	US-10-348-052-20	Sequence 20, Appl
44	301	10.0	524	16	US-10-622-011-20	Sequence 20, Appl
45	301	10.0	907	15	US-10-348-052-29	Sequence 29, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-969-896-11  
; Sequence 11, Application US/09969896  
; Publication No. US20030125533A1  
; GENERAL INFORMATION:  
; APPLICANT: Kossida, Sophia  
; TITLE OF INVENTION: Regulation of human Sphingosine  
; TITLE OF INVENTION: Kinase-Like Protein  
; FILE REFERENCE: 004974.00594  
; CURRENT APPLICATION NUMBER: US/09/969,896  
; PRIOR FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: US 60/238,005  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: US 60/314,113  
; PRIOR FILING DATE: 2001-08-23  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 562  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-969-896-11

Query Match 100.0%; Score 3025; DB 10; Length 562;  
Best Local Similarity 100.0%; Pred. No. 5.2e-281;  
Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 HEANGAPAGLGVAPPAAWRTSPAEMGATGAAPLQSVLVWVKQRCVLSLPPARALLRW 60

Db 1 HEANGAPAGLGVAPPAAWRTSPAEMGATGAAPLQSVLVWVKQRCVLSLPPARALLRW 60

Oy 61 RSPGPGAGACADACSPVSEIIAIVETDVHGKQSGKQMKPKPAFTVHCVRARRH 120

Db 61 RSPGPGAGACADACSPVSEIIAIVETDVHGKQSGKQMKPKPAFTVHCVRARRH 120



```
Qy 266 TDCVYSTVGTSDAETSAHIIWVGDLSAMDVSSVHHNSTLLRYSVSLGCGYFGDIIKDS 325
Db 241 TDCVYSTVGTSDAETSAHIIWVGDLSAMDVSSVHHNSTLLRYSVSLGCGYFGDIIKDS 300
Qy 326 EKKRWGLGARYDFSGKTLFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQ 385
Db 301 EKKRWGLGARYDFSGKTLFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQ 360
Qy 386 LEEQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSGL 445
Db 361 LEEQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSGL 420
Qy 446 ILIRKCSRFNLFRLIRHTNQDQDFTFVEVYRVKKQFTSKHMEDESDLKEGGKRF 505
Db 421 ILIRKCSRFNLFRLIRHTNQDQDFTFVEVYRVKKQFTSKHMEDESDLKEGGKRF 480
Qy 506 GHICSHSPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDHS 562
Db 481 GHICSHSPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDHS 537

RESULT 4
US-10-262-511-40
; Sequence 40, Application US/10262511
; Publication No. US20040038223A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; PRIOR FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
```

```
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 40
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-511-40

Query Match 95.5%; Score 2888; DB 15; Length 537;
Best Local Similarity 100.0%; Pred. No. 6,9e-268;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 MGATGAAPLQSVLWVQKQRCVSLPARALLRWRSFPGAGAPGADACSPVSEIIAV 85
Db 1 MGATGAAPLQSVLWVQKQRCVSLPARALLRWRSFPGAGAPGADACSPVSEIIAV 60

Qy 86 EETDVHGKHGSGKQWQKEXPYAFTVHCVRARHRWKWAQVTFWCPEEQIHLWLQTLR 145
Db 61 EETDVHGKHGSGKQWQKEXPYAFTVHCVRARHRWKWAQVTFWCPEEQIHLWLQTLR 120

Qy 146 EMLEKLSRPHLLVFINPFGGKGQKRIYERKVAPLFTLASITTDIIIVTSHANOAKETL 205
Db 121 EMLEKLSRPHLLVFINPFGGKGQKRIYERKVAPLFTLASITTDIIIVTSHANOAKETL 180

Qy 206 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRVAVLPSSLRIGIIPAGS 265
Db 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRVAVLPSSLRIGIIPAGS 240

Qy 266 TDCVYSTVGTSDAETSAHIIWVGDLSAMDVSSVHHNSTLLRYSVSLGCGYFGDIIKDS 325
Db 241 TDCVYSTVGTSDAETSAHIIWVGDLSAMDVSSVHHNSTLLRYSVSLGCGYFGDIIKDS 300

Qy 326 EKKRWGLGARYDFSGKTLFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQ 385
Db 301 EKKRWGLGARYDFSGKTLFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQ 360

Qy 386 LEEQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSGL 445
Db 361 LEEQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSGL 420

Qy 446 ILIRKCSRFNLFRLIRHTNQDQDFTFVEVYRVKKQFTSKHMEDESDLKEGGKRF 505
Db 421 ILIRKCSRFNLFRLIRHTNQDQDFTFVEVYRVKKQFTSKHMEDESDLKEGGKRF 480

Qy 506 GHICSHSPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDHS 562
Db 481 GHICSHSPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDHS 537

RESULT 5
US-10-631-958-10
; Sequence 10, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; PRIOR FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
```

		PRIORITY FILING DATE: 2001-08-23		Best Local Similarity 100.0%; Pred. No. 6.9e-268;		Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
		; NUMBER OF SEQ ID NOS: 16					
		; SOFTWARE: FastSeq for Windows Version 4.0					
		; SEQ ID NO 10					
		; LENGTH: 537					
		; TYPE: PRT					
		; ORGANISM: Homo sapiens					
		US-10-631-958-10					
		Query Match 95.5%; Score 2888; DB 16; Length 537;					
		Best Local Similarity 100.0%; Pred. No. 6.9e-268;					
		Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	26	MGATGAAPLQSVLWVKQRCVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV	85	Qy	26	MGATGAAPLQSVLWVKQRCVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV	85
Db	1	MGATGAAPLQSVLWVKQRCVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV	60	Db	1	MGATGAAPLQSVLWVKQRCVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV	60
Qy	86	EETDVHGKGSGKWKQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR	145	Qy	86	EETDVHGKGSGKWKQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR	145
Db	61	EETDVHGKGSGKWKQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR	120	Db	61	EETDVHGKGSGKWKQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR	120
Qy	146	EMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL	205	Qy	146	EMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL	205
Db	121	EMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL	180	Db	121	EMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL	180
Qy	206	YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGS	265	Qy	206	YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGS	265
Db	181	YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGS	240	Db	181	YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGS	240
Qy	266	TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS	325	Qy	266	TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS	325
Db	241	TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS	300	Db	241	TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS	300
Qy	326	EKKRWGLIARYDFSGIKTFLSHHCYEGTGSFLPAQHTVGSPRDRKPCRACGFCVCRQSKQ	385	Qy	326	EKKRWGLIARYDFSGIKTFLSHHCYEGTGSFLPAQHTVGSPRDRKPCRACGFCVCRQSKQ	385
Db	301	EKKRWGLIARYDFSGIKTFLSHHCYEGTGSFLPAQHTVGSPRDRKPCRACGFCVCRQSKQ	360	Db	301	EKKRWGLIARYDFSGIKTFLSHHCYEGTGSFLPAQHTVGSPRDRKPCRACGFCVCRQSKQ	360
Qy	386	LEBEQKALYGLEAAEDVEEQQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL	445	Qy	386	LEBEQKALYGLEAAEDVEEQQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL	445
Db	361	LEBEQKALYGLEAAEDVEEQQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL	420	Db	361	LEBEQKALYGLEAAEDVEEQQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL	420
Qy	446	ILIRKCSRNFRLRIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGGKKRF	505	Qy	446	ILIRKCSRNFRLRIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGGKKRF	505
Db	421	ILIRKCSRNFRLRIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGGKKRF	480	Db	421	ILIRKCSRNFRLRIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGGKKRF	480
Qy	506	GHICSSHPSCCCTVSNSSMNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSDS	562	Qy	506	GHICSSHPSCCCTVSNSSMNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSDS	562
Db	481	GHICSSHPSCCCTVSNSSMNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSDS	537	Db	481	GHICSSHPSCCCTVSNSSMNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSDS	537
		RESULT 7					
		US-10-315-597A-2					
		; Sequence 2, Application US/10315597A					
		; Publication No. US20030162206A1					
		; GENERAL INFORMATION:					
		; APPLICANT: Sugiyura, Masako					
		; APPLICANT: Kono, Keita					
		; TITLE OF INVENTION: Ceramide Kinase and DNA Encoding It					
		; FILE REFERENCE: 02658CIP/HG					
		; CURRENT APPLICATION NUMBER: US/10/315,597A					
		; CURRENT FILING DATE: 2002-12-10					
		; PRIOR APPLICATION NUMBER: JP 2000-178039					
		; PRIOR FILING DATE: 2000-06-14					
		; NUMBER OF SEQ ID NOS: 4					
		; SOFTWARE: PatentIn Ver. 2.0					
		; SEQ ID NO 2					
		; LENGTH: 537					
		; TYPE: PRT					
		; ORGANISM: Homo sapiens					
		US-10-315-597A-2					
		Query Match 95.2%; Score 2880; DB 14; Length 537;					
		Best Local Similarity 99.6%; Pred. No. 4.1e-267;					
		Matches 535; Conservative 1; Mismatches 1; Indels 0; Gaps 0;					
Qy	26	MGATGAAPLQSVLWVKQRCVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV	85	Qy	26	MGATGAAPLQSVLWVKQRCVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV	85
Db	1	MGATGAAPLQSVLWVKQRCVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV	60	Db	1	MGATGAAPLQSVLWVKQRCVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV	60
Qy	86	EETDVHGKGSGKWKQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR	145	Qy	86	EETDVHGKGSGKWKQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR	145
Db	61	EETDVHGKGSGKWKQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR	120	Db	61	EETDVHGKGSGKWKQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR	120
		Query Match 95.5%; Score 2888; DB 16; Length 537;					
		Best Local Similarity 100.0%; Pred. No. 6.9e-268;					
		Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	26	MGATGAAPLQSVLWVKQRCVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV	85	Qy	26	MGATGAAPLQSVLWVKQRCVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV	85
Db	1	MGATGAAPLQSVLWVKQRCVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV	60	Db	1	MGATGAAPLQSVLWVKQRCVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV	60
Qy	86	EETDVHGKGSGKWKQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR	145	Qy	86	EETDVHGKGSGKWKQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR	145
Db	61	EETDVHGKGSGKWKQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR	120	Db	61	EETDVHGKGSGKWKQMEKPYAFTVHCVKRARRHRWKWAQVTFWCPPEEQCHLWLQTLR	120
Qy	146	EMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL	205	Qy	146	EMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL	205
Db	121	EMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL	180	Db	121	EMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL	180
Qy	206	YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGS	265	Qy	206	YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGS	265
Db	181	YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGS	240	Db	181	YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGS	240
Qy	266	TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS	325	Qy	266	TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS	325
Db	241	TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS	300	Db	241	TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS	300
Qy	326	EKKRWGLIARYDFSGIKTFLSHHCYEGTGSFLPAQHTVGSPRDRKPCRACGFCVCRQSKQ	385	Qy	326	EKKRWGLIARYDFSGIKTFLSHHCYEGTGSFLPAQHTVGSPRDRKPCRACGFCVCRQSKQ	385
Db	301	EKKRWGLIARYDFSGIKTFLSHHCYEGTGSFLPAQHTVGSPRDRKPCRACGFCVCRQSKQ	360	Db	301	EKKRWGLIARYDFSGIKTFLSHHCYEGTGSFLPAQHTVGSPRDRKPCRACGFCVCRQSKQ	360
Qy	386	LEBEQKALYGLEAAEDVEEQQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL	445	Qy	386	LEBEQKALYGLEAAEDVEEQQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL	445
Db	361	LEBEQKALYGLEAAEDVEEQQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL	420	Db	361	LEBEQKALYGLEAAEDVEEQQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL	420
Qy	446	ILIRKCSRNFRLRIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGGKKRF	505	Qy	446	ILIRKCSRNFRLRIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGGKKRF	505
Db	421	ILIRKCSRNFRLRIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGGKKRF	480	Db	421	ILIRKCSRNFRLRIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGGKKRF	480
Qy	506	GHICSSHPSCCCTVSNSSMNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSDS	562	Qy	506	GHICSSHPSCCCTVSNSSMNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSDS	562
Db	481	GHICSSHPSCCCTVSNSSMNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSDS	537	Db	481	GHICSSHPSCCCTVSNSSMNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSDS	537
		RESULT 6					
		US-10-618-941-121					
		; Sequence 121, Application US/10618941					
		; Publication No. US2004019792A1					
		; GENERAL INFORMATION:					
		; APPLICANT: WHYTE, DAVID					
		; APPLICANT: MANNING, GERARD					
		; APPLICANT: CAENEPEEL, SEAN					
		; TITLE OF INVENTION: NOVEL KINASES					
		; FILE REFERENCE: 034536-0321					
		; CURRENT APPLICATION NUMBER: US/10/618,941					
		; CURRENT FILING DATE: 2003-07-15					
		; PRIOR APPLICATION NUMBER: 60/395,632					
		; PRIOR FILING DATE: 2002-07-15					
		; NUMBER OF SEQ ID NOS: 143					
		; SOFTWARE: PatentIn version 3.2					
		; SEQ ID NO 121					
		; LENGTH: 537					
		; TYPE: PRT					
		; ORGANISM: Homo sapiens					
		US-10-618-941-121					
		Query Match 95.5%; Score 2888; DB 16; Length 537;					



```
QY 146 EMLEKLTSPKHLVFINPFGKGQGGKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 205
DB 121 EMLEKLTSPKHLVFINPFGKGQGGKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180
QY 206 YEINIDKYDGIIVCGDGMFSEVLHGLIGRTORSAGVDQNHPRVAVLPSSLRIGIIPAGS 265
DB 181 YEINIDKYDGIIVCGDGMFSEVLHGLIGRTORSAGVDQNHPRVAVLPSSLRIGIIPAGP 240
QY 266 TDCVYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGIFYGFIIDKXS 325
DB 241 TDCVYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGIFYGFIIDKXS 300
QY 326 EKKRWLGILARYDFSLGKTLFSLHHCYEGTFSFLPAQHTVGSPRDRKPCRAGFCVCRSQKQ 385
DB 301 EKKRWLGILARYDFSLGKTLFSLHHCYEGTFSFLPAQHTVGSPRDRKPCRAGFCVCRSQKQ 360
QY 386 LEEEOKKALYGLEAAEDVEEWOVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 445
DB 361 LEEEOKKALYGLEAAEDVEEWOVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 420
QY 446 ILIRKCSRFNLFIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLKEGGKKRF 505
DB 421 ILIRKCSRFNLFIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLKEGGKKRF 480
QY 506 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 562
DB 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 537

RESULT 8
US-09-784-810A-6
; Sequence 6, Application US/09784810A
; Patent No. US20020082203A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: SAME
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/09/784,810A
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-784-810A-6

Query Match 81.2%; Score 2456.5; DB 9; Length 471;
Best Local Similarity 97.5%; Pred. No. 1.6e-226;
Matches 459; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

QY 103 MEKPYAFTVHCVRARRHRKWAQVTFWCPPEQLCHLWLQTLREMLEKLTSPKHLVFI 162
DB 1 MEKPYAFTVHCVRARRHRKWAQVTFWCPPEQLCHLWLQTLREMLEKLTSPKHLVFI 60

QY 163 NPFQGGQGGKRIYERKVAFLFTLASITTDII-----VTEHANOAKETLYEINID 211
DB 61 NPFQGGQGGKRIYERKVAFLFTLASITTDIIIGNKFVYVYVEVITEHANOAKETLYEINID 120

QY 212 KYDGIIVCGDGMFSEVLHGLIGRTORSAGVDQNHPRVAVLPSSLRIGIIPAGSTDVCY 271
DB 121 KYDGIIVCGDGMFSEVLHGLIGRTORSAGVDQNHPRVAVLPSSLRIGIIPAGSTDVCY 180

QY 272 STVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGIFYGFIIDKSEKKRWL 331
DB 181 STVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGIFYGFIIDKSEKKRWL 240

QY 332 GLARYDFSLGKTLFSLHHCYEGTFSFLPAQHTVGSPRDRKPCRAGFCVCRSQKQLEEBQK 391
DB 241 GLARYDFSLGKTLFSLHHCYEGTFSFLPAQHTVGSPRDRKPCRAGFCVCRSQKQLEEBQK 300

QY 392 KALYGLEAAEDVEEWOVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDLIRKRC 451
DB 301 KALYGLEAAEDVEEWOVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDLIRKRC 360

QY 452 SRNFNLFRLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLKEGGKKRFHICSS 511
DB 361 SRNFNLFRLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLKEGGKKRFHICSS 420
```

```
QY 332 GLARYDFSLGKTLFSLHHCYEGTFSFLPAQHTVGSPRDRKPCRAGFCVCRSQKQLEEBQK 391
DB 241 GLARYDFSLGKTLFSLHHCYEGTFSFLPAQHTVGSPRDRKPCRAGFCVCRSQKQLEEBQK 300
QY 392 KALYGLEAAEDVEEWOVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDLIRKRC 451
DB 301 KALYGLEAAEDVEEWOVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDLIRKRC 360
QY 452 SRNFNLFRLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLKEGGKKRFHICSS 511
DB 361 SRNFNLFRLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLKEGGKKRFHICSS 420
QY 512 HPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 562
DB 421 HPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 471

RESULT 9
US-10-876-281-6
; Sequence 6, Application US/10876281
; Publication No. US20050123942A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: SAME
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/10/876,281
; CURRENT FILING DATE: 2004-06-24
; PRIOR APPLICATION NUMBER: US/09/784,810
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-876-281-6

Query Match 81.2%; Score 2456.5; DB 18; Length 471;
Best Local Similarity 97.5%; Pred. No. 1.6e-226;
Matches 459; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

QY 103 MEKPYAFTVHCVRARRHRKWAQVTFWCPPEQLCHLWLQTLREMLEKLTSPKHLVFI 162
DB 1 MEKPYAFTVHCVRARRHRKWAQVTFWCPPEQLCHLWLQTLREMLEKLTSPKHLVFI 60

QY 163 NPFQGGQGGKRIYERKVAFLFTLASITTDII-----VTEHANOAKETLYEINID 211
DB 61 NPFQGGQGGKRIYERKVAFLFTLASITTDIIIGNKFVYVYVEVITEHANOAKETLYEINID 120

QY 212 KYDGIIVCGDGMFSEVLHGLIGRTORSAGVDQNHPRVAVLPSSLRIGIIPAGSTDVCY 271
DB 121 KYDGIIVCGDGMFSEVLHGLIGRTORSAGVDQNHPRVAVLPSSLRIGIIPAGSTDVCY 180

QY 272 STVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGIFYGFIIDKSEKKRWL 331
DB 181 STVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGIFYGFIIDKSEKKRWL 240

QY 332 GLARYDFSLGKTLFSLHHCYEGTFSFLPAQHTVGSPRDRKPCRAGFCVCRSQKQLEEBQK 391
DB 241 GLARYDFSLGKTLFSLHHCYEGTFSFLPAQHTVGSPRDRKPCRAGFCVCRSQKQLEEBQK 300

QY 392 KALYGLEAAEDVEEWOVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDLIRKRC 451
DB 301 KALYGLEAAEDVEEWOVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDLIRKRC 360

QY 452 SRNFNLFRLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLKEGGKKRFHICSS 511
DB 361 SRNFNLFRLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLKEGGKKRFHICSS 420
```

```
Qy 512 HPSCCCTVSNSSMNCDEVLHSPAIEVRVHCOLVRLFARGTEENPKPDSSH 562
Db 421 HPSCCCTVSNSSMNCDEVLHSPAIEVRVHCOLVRLFARGTEENPKPDSSH 471

RESULT 10
US-09-784-810A-11
; Sequence 11, Application US/09784810A
; Patent No. US20020082203A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; FILE OF INVENTION: SAME
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/09/784,810A
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-784-810A-11

Query Match 54.2%; Score 1640.5; DB 9; Length 326;
Best Local Similarity 96.3%; Pred. No. 2.le-148;
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

Qy 155 PKHLLVFNPFGKGQKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 203
Db 1 PKHLLVFNPFGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVNVVEVITEHANOAKE 60

Qy 204 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPA 263
Db 61 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPA 120

Qy 264 GSTDCVCYSTVGTSDAETSALHIVVGDSDLAMDVSSVHHNSTLLRYSVSLGFGYGIK 323
Db 121 GSTDCVCYSTVGTSDAETSALHIVVGDSDLAMDVSSVHHNSTLLRYSVSLGFGYGIK 180

Qy 324 DSEKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCQSK 383
Db 181 DSEKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCQSK 240

Qy 384 QOLEEEQKALYGLEAAEDVEWQVCGKFLAINATNMSCACRSPRGLSPAHLGDGSS 443
Db 241 QOLEEEQKALYGLEAAEDVEWQVCGKFLAINATNMSCACRSPRGLSPAHLGDGSS 300

Qy 444 DLILIRKCSRFNRLFLIRHTNQDQ 469
Db 301 DLILIRKCSRFNRLFLIRHTNQDQ 326

RESULT 11
US-09-969-896-2
; Sequence 2, Application US/09969896
; Publication No. US20030125533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; ;
```

```
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-896-2

Query Match 54.2%; Score 1640.5; DB 10; Length 326;
Best Local Similarity 96.3%; Pred. No. 2.le-148;
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

Qy 155 PKHLLVFNPFGKGQKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 203
Db 1 PKHLLVFNPFGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVNVVEVITEHANOAKE 60

Qy 204 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPA 263
Db 61 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPA 120

Qy 264 GSTDCVCYSTVGTSDAETSALHIVVGDSDLAMDVSSVHHNSTLLRYSVSLGFGYGIK 323
Db 121 GSTDCVCYSTVGTSDAETSALHIVVGDSDLAMDVSSVHHNSTLLRYSVSLGFGYGIK 180

Qy 324 DSEKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCQSK 383
Db 181 DSEKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCQSK 240

Qy 384 QOLEEEQKALYGLEAAEDVEWQVCGKFLAINATNMSCACRSPRGLSPAHLGDGSS 443
Db 241 QOLEEEQKALYGLEAAEDVEWQVCGKFLAINATNMSCACRSPRGLSPAHLGDGSS 300

Qy 444 DLILIRKCSRFNRLFLIRHTNQDQ 469
Db 301 DLILIRKCSRFNRLFLIRHTNQDQ 326

RESULT 12
US-10-631-958-2
; Sequence 2, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-631-958-2

Query Match 54.2%; Score 1640.5; DB 16; Length 326;
Best Local Similarity 96.3%; Pred. No. 2.le-148;
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

Qy 155 PKHLLVFNPFGKGQKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 203
Db 1 PKHLLVFNPFGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVNVVEVITEHANOAKE 60

Qy 204 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPA 263
Db 61 TLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPA 120

Qy 264 GSTDCVCYSTVGTSDAETSALHIVVGDSDLAMDVSSVHHNSTLLRYSVSLGFGYGIK 323
Db 121 GSTDCVCYSTVGTSDAETSALHIVVGDSDLAMDVSSVHHNSTLLRYSVSLGFGYGIK 180

Qy 324 DSEKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCQSK 383
Db 181 DSEKRWLGLARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCQSK 240

Qy 384 QOLEEEQKALYGLEAAEDVEWQVCGKFLAINATNMSCACRSPRGLSPAHLGDGSS 443
Db 241 QOLEEEQKALYGLEAAEDVEWQVCGKFLAINATNMSCACRSPRGLSPAHLGDGSS 300

Qy 444 DLILIRKCSRFNRLFLIRHTNQDQ 469
Db 301 DLILIRKCSRFNRLFLIRHTNQDQ 326
```

QY 264 GSTDCVCTVGTSDAETSALHIVVGDLSLAMDVSVVHNSLTLRYSVSLLGFGFYGDIK 323  
DB |||||  
121 GSTDCVCTVGTSDAETSALHIVVGDLSLAMDVSVVHNSLTLRYSVSLLGFGFYGDIK 180  
QY 324 DSEKRWLGARYDFSGLKTFLSHHCYEGTYSFLPAQHTVGSPPDRKPCACGFCVCRSK 383  
DB |||||  
181 DSEKRWLGARYDFSGLKTFLSHHCYEGTYSFLPAQHTVGSPPDRKPCACGFCVCRSK 240  
QY 384 QOLEEQKALYGLEAAEDVEWQVCGKFLAINATNMSCARRSPRGLSPAHLGDSS 443  
DB |||||  
241 QOLEEQKALYGLEAAEDVEWQVCGKFLAINATNMSCARRSPRGLSPAHLGDSS 300  
QY 444 DLILIRKSRNFLRFLIRHTNQDQ 469  
DB |||||  
301 DLILIRKSRNFLRFLIRHTNQDQ 326

## RESULT 13

US-10-876-281-11  
; Sequence 11, Application US/10876281  
; Publication No. US20050123942A1  
; GENERAL INFORMATION:  
; APPLICANT: RASTELLI, LUCA  
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING  
; FILE REFERENCE: 10716-08  
; CURRENT APPLICATION NUMBER: US/10/876,281  
; PRIOR FILING DATE: 2004-06-24  
; PRIOR APPLICATION NUMBER: US/09/784,810  
; PRIOR FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: 60/182,360  
; PRIOR FILING DATE: 2000-02-14  
; PRIOR APPLICATION NUMBER: 60/191,261  
; PRIOR FILING DATE: 2000-03-22  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 326  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-876-281-11

Query Match 54.2%; Score 1640.5; DB 18; Length 326;

Best Local Similarity 96.3%; Pred. No. 2.1e-148;

Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

QY 155 PKHLVFNPGGKGQKRIYERKVAPLFTLASITTDII-----VTEHANOAKE 203  
DB |||||  
1 PKHLVFNPGGKGQKRIYERKVAPLFTLASITTDIIIGNKFYVYVEVITEHANOAKE 60  
QY 204 TLYEINIDKYDGVCGGDMFSEVLHGLIGRTORSAGVDQNHPRVLPVPSLRIGIIPA 263  
DB |||||  
61 TLYEINIDKYDGVCGGDMFSEVLHGLIGRTORSAGVDQNHPRVLPVPSLRIGIIPA 120  
QY 264 GSTDCVCTVGTSDAETSALHIVVGDLSLAMDVSVVHNSLTLRYSVSLLGFGFYGDIK 323  
DB |||||  
121 GSTDCVCTVGTSDAETSALHIVVGDLSLAMDVSVVHNSLTLRYSVSLLGFGFYGDIK 180  
QY 324 DSEKRWLGARYDFSGLKTFLSHHCYEGTYSFLPAQHTVGSPPDRKPCACGFCVCRSK 383  
DB |||||  
181 DSEKRWLGARYDFSGLKTFLSHHCYEGTYSFLPAQHTVGSPPDRKPCACGFCVCRSK 240  
QY 384 QOLEEQKALYGLEAAEDVEWQVCGKFLAINATNMSCARRSPRGLSPAHLGDSS 443  
DB |||||  
241 QOLEEQKALYGLEAAEDVEWQVCGKFLAINATNMSCARRSPRGLSPAHLGDSS 300  
QY 444 DLILIRKSRNFLRFLIRHTNQDQ 469  
DB |||||  
301 DLILIRKSRNFLRFLIRHTNQDQ 326

## RESULT 14

Query Match 32.6%; Score 986; DB 18; Length 182;

US-09-784-810A-29  
; Sequence 29, Application US/09784810A  
; Patent No. US20020082203A1  
; GENERAL INFORMATION:  
; APPLICANT: RASTELLI, LUCA  
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING  
; FILE REFERENCE: 10716-08  
; CURRENT APPLICATION NUMBER: US/09/784,810A  
; PRIOR FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: 60/182,360  
; PRIOR FILING DATE: 2000-02-14  
; PRIOR APPLICATION NUMBER: 60/191,261  
; PRIOR FILING DATE: 2000-03-22  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 29  
; LENGTH: 182  
; TYPE: PRT  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: 80432911  
US-09-784-810A-29

Query Match 32.6%; Score 986; DB 9; Length 182;

Best Local Similarity 100.0%; Pred. No. 6.4e-86;

Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 AGAFGADACSVPVSEIIAVEETDVHGKHQSGKQKMEKPYAFTVHCVKRARRHKKWAQ 126  
DB |||||  
1 AGAFGADACSVPVSEIIAVEETDVHGKHQSGKQKMEKPYAFTVHCVKRARRHKKWAQ 60  
QY 127 VTFWCPPEQLCHLWLQTLREMLEKLTSRPKLLVFINPFGGKGQKRIYERKVAPLFTLA 186  
DB |||||  
61 VTFWCPPEQLCHLWLQTLREMLEKLTSRPKLLVFINPFGGKGQKRIYERKVAPLFTLA 120  
QY 187 SITTDIIIVTEHANOAKETLYEINIDKYDGVCGGDMFSEVLHGLIGRTORSAGVDQNH 246  
DB |||||  
121 SITTDIIIVTEHANOAKETLYEINIDKYDGVCGGDMFSEVLHGLIGRTORSAGVDQNH 180  
QY 247 PR 248  
DB ||| PR 182

## RESULT 15

US-10-876-281-29  
; Sequence 29, Application US/10876281  
; Publication No. US20050123942A1  
; GENERAL INFORMATION:  
; APPLICANT: RASTELLI, LUCA  
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING  
; FILE REFERENCE: 10716-08  
; CURRENT APPLICATION NUMBER: US/10/876,281  
; PRIOR FILING DATE: 2004-06-24  
; PRIOR APPLICATION NUMBER: US/09/784,810  
; PRIOR FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: 60/182,360  
; PRIOR FILING DATE: 2000-02-14  
; PRIOR APPLICATION NUMBER: 60/191,261  
; PRIOR FILING DATE: 2000-03-22  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 29  
; LENGTH: 182  
; TYPE: PRT  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: 80432911  
US-10-876-281-29

Best Local Similarity 100.0%; Pred. No. 6.4e-86;  
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	67	AGAPGADACSVPVSEIIIAVEETDVHGKHQSGKWQMEKPYAFTVHCVGARRHRWKWQAQ	126
Db	1	AGAPGADACSVPVSEIIIAVEETDVHGKHQSGKWQMEKPYAFTVHCVGARRHRWKWQAQ	60
Qy	127	VTFWCPBEEQLCHLWLQTLREMLEKLTSPKHLVFINPFGGKGQKRIYERKVAPLFTLA	186
Db	61	VTFWCPBEEQLCHLWLQTLREMLEKLTSPKHLVFINPFGGKGQKRIYERKVAPLFTLA	120
Qy	187	SITTDIIIVTEHANOAKETLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNH	246
Db	121	SITTDIIIVTEHANOAKETLYEINIDKYDGI VCVGGDGMFSEVLHGLIGRTORSAGVDQNH	180
Qy	247	PR 248	
Db	181	PR 182	

Search completed: September 3, 2005, 04:32:36  
Job time : 112.217 secs

GenCore version 5.1.1.6  
Copyright (C) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 04:08:02; Search time 106.269 Seconds  
(without alignments)  
1990.064 Million cell updates/sec

Title: US-10-631-958-10

Perfect score: 2888

Sequence: 1 MGATGAAPLQSLVWVQKQR.....QLVRLFARGIENPDPDHS 537

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1774312 seqs, 393823214 residues

Total number of hits satisfying chosen parameters: 1774312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17:	/cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
19:	/cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20:	/cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2888	100.0	537	10	US-09-969-896-10
2	2888	100.0	537	15	US-10-262-511-40
3	2888	100.0	537	16	US-10-631-958-10
4	2888	100.0	537	16	US-10-618-941-121
5	2888	100.0	562	10	US-09-969-896-11
6	2888	100.0	562	16	US-10-631-958-11
7	2880	99.7	537	14	US-10-315-597A-2
8	2456.5	85.1	471	9	US-09-784-810A-6
9	2456.5	85.1	471	18	US-10-876-281-6
10	1640.5	56.8	326	9	US-09-784-810A-11
11	1640.5	56.8	326	10	US-09-969-896-2

12	1640.5	56.8	326	16	US-10-631-958-2	Sequence 2, Appl
13	1640.5	56.8	326	18	US-10-876-281-11	Sequence 11, Appl
14	986	34.1	182	9	US-09-784-810A-29	Sequence 29, Appl
15	986	34.1	182	18	US-10-876-281-29	Sequence 8, Appl
16	647	22.4	144	9	US-09-784-810A-8	Sequence 8, Appl
17	647	22.4	144	18	US-10-876-281-8	Sequence 361076,
18	611.5	21.2	588	16	US-10-425-115-361076	Sequence 270874,
19	597.5	20.7	575	16	US-10-425-115-270874	Sequence 195284,
20	480.5	16.6	613	16	US-10-437-963-195284	Sequence 133133,
21	399.5	13.8	484	16	US-10-437-963-133133	Sequence 190238,
22	389	13.5	487	15	US-10-424-599-190238	Sequence 213665,
23	362.5	12.6	760	16	US-10-425-115-213665	Sequence 249553,
24	357.5	12.4	490	16	US-10-425-115-249553	Sequence 249555,
25	344.5	11.9	381	16	US-10-425-115-249555	Sequence 202261,
26	341	11.8	520	15	US-10-424-599-202261	Sequence 10, Appl
27	335	11.6	79	9	US-09-784-810A-10	Sequence 10, Appl
28	335	11.6	79	18	US-10-876-281-10	Sequence 53957, A
29	334	11.6	245	15	US-10-425-114-53957	Sequence 17631,
30	331.5	11.5	446	16	US-10-437-963-197631	Sequence 180379,
31	330.5	11.4	757	16	US-10-437-963-180379	Sequence 282512,
32	326.5	11.3	470	16	US-10-425-115-282512	Sequence 68423, A
33	326.5	11.3	521	15	US-10-425-114-68423	Sequence 264087,
34	313	10.8	677	16	US-10-425-115-264087	Sequence 44927, A
35	312.5	10.8	373	16	US-10-767-701-44927	Sequence 161184,
36	312.5	10.8	552	16	US-10-437-963-161184	Sequence 28, Appl
37	310.5	10.8	641	15	US-10-348-052-28	Sequence 28, Appl
38	310.5	10.8	641	16	US-10-622-011-28	Sequence 19, Appl
39	306	10.6	490	14	US-10-053-510-19	Sequence 19, Appl
40	306	10.6	490	15	US-10-348-052-19	Sequence 19, Appl
41	306	10.6	490	16	US-10-622-011-19	Sequence 20, Appl
42	301	10.4	524	14	US-10-053-510-20	Sequence 20, Appl
43	301	10.4	524	15	US-10-348-052-20	Sequence 20, Appl
44	301	10.4	524	16	US-10-622-011-20	Sequence 25, Appl
45	301	10.4	907	15	US-10-348-052-29	

#### ALIGNMENTS

RESULT 1  
US-09-969-896-10  
; Sequence 10, Application US/09969896  
; Publication No. US2003012553A1  
; GENERAL INFORMATION:  
; APPLICANT: Kossida, Sophia  
; TITLE OF INVENTION: Regulation of human Sphingosine  
; FILE OF INVENTION: Kinase-Like Protein  
; FILE REFERENCE: 004974.00594  
; CURRENT APPLICATION NUMBER: US/09/969,896  
; CURRENT FILING DATE: 2001-10-04  
; PRIOR FILING DATE: 2001-10-06  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR FILING DATE: 2001-08-23  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 537  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-969-896-10

Query Match 100.0%; Score 2888; DB 10; Length 537;  
Best Local Similarity 100.0%; Pred. No. 3e-278;  
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGATGAAPLQSLVWVQKQRCAVSLPARALLRWRSPGCGAGAPGADACSVPVSEIAV 60

Db 1 MGATGAAPLQSLVWVQKQRCAVSLPARALLRWRSPGCGAGAPGADACSVPVSEIAV 60

Qy 61 EETDVHGKQSGKWKQKXPYAFVHCVKRHRWRKWAQVTFWCPEQLCHLWLTLR 120

Db 61 EETDVHGKQSGKWKQKXPYAFVHCVKRHRWRKWAQVTFWCPEQLCHLWLTLR 120

Qy 121 EMLEKLTSPKHLVFINPFGKGQGGKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180  
Db 121 EMLEKLTSPKHLVFINPFGKGQGGKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180  
Qy 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGS 240  
Db 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGS 240  
Qy 241 TDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGCGFYGDIIKDS 300  
Db 241 TDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGCGFYGDIIKDS 300  
Qy 301 EKKRWGLGARYDFSGLKTFLSHHCVEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQ 360  
Db 301 EKKRWGLGARYDFSGLKTFLSHHCVEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQ 360  
Qy 361 LEEQKALYGLEAAEDVEEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420  
Db 361 LEEQKALYGLEAAEDVEEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420  
Qy 421 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLKEGGKKRF 480  
Db 421 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLKEGGKKRF 480  
Qy 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 537  
Db 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 537

RESULT 2

US-10-262-511-40  
; Sequence 40, Application US/10262511  
; Publication No. US20040038223A1  
; GENERAL INFORMATION:  
; APPLICANT: Smithson, Glennda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Catterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shimkets, Richard A.  
; APPLICANT: Rothenberg, Mark E.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Ages, Michele L.  
; APPLICANT: Berghs, Constance  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-462C  
; CURRENT APPLICATION NUMBER: US/10/262,511  
; PRIORITY FILING DATE: 2003-05-28  
; PRIOR APPLICATION NUMBER: 60/326,483  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/373,815  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,917

; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/381,642  
; PRIOR FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: 60/328,029  
; PRIOR FILING DATE: 2002-10-09  
; PRIOR APPLICATION NUMBER: 60/381,038  
; PRIOR FILING DATE: 2002-05-16  
; PRIOR APPLICATION NUMBER: 60/328,056  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/373,260  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/373,826  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,435  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: Curaseq1ist version 0.1  
; SEQ ID NO 40  
; LENGTH: 537  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-262-511-40  
  
Query Match 100.0%; Score 2888; DB 15; Length 537;  
Best Local Similarity 100.0%; Pred. No. 3e-278;  
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 MGATGAAEPLQSVLWVKQRCVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV 60  
Db 1 MGATGAAEPLQSVLWVKQRCVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV 60  
Qy 61 BETDVGKHGSGKWKQWKPKYAFTHVCVKRARRHRWKQAQVTFWCPEEOLCHLWLQTLR 120  
Db 61 BETDVGKHGSGKWKQWKPKYAFTHVCVKRARRHRWKQAQVTFWCPEEOLCHLWLQTLR 120  
Qy 121 EMLEKLTSPKHLVFINPFGKGQGGKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180  
Db 121 EMLEKLTSPKHLVFINPFGKGQGGKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180  
Qy 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGS 240  
Db 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGS 240  
Qy 241 TDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGCGFYGDIIKDS 300  
Db 241 TDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGCGFYGDIIKDS 300  
Qy 301 EKKRWGLGARYDFSGLKTFLSHHCVEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQ 360  
Db 301 EKKRWGLGARYDFSGLKTFLSHHCVEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQ 360  
Qy 361 LEEQKALYGLEAAEDVEEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420  
Db 361 LEEQKALYGLEAAEDVEEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420  
Qy 421 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLKEGGKKRF 480  
Db 421 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLKEGGKKRF 480  
Qy 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 537  
Db 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 537  
  
RESULT 3  
US-10-631-958-10  
; Sequence 10, Application US/10631958  
; Publication No. US20040192580A1  
; GENERAL INFORMATION:  
; APPLICANT: Kossida, Sophia  
; TITLE OF INVENTION: Regulation of human Sphingosine  
; TITLE OF INVENTION: Kinase-Like Protein

```
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; PRIOR FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 537
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-631-958-10

Query Match      100.0%; Score 2888; DB 16; Length 537;
Best Local Similarity 100.0%; Pred. No. 3e-278;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGATGAAPLQSLVWVKQRCVSLPARALLRWRSFPGAGAPGADACSVPVSEIIAV 60
Db 1 MGATGAAPLQSLVWVKQRCVSLPARALLRWRSFPGAGAPGADACSVPVSEIIAV 60

Qy 61 BETDVHGKHQSGKWKQMEKPYAFTVHCVRARRHRWKWAQVTFWCPBEQCHLWLQTLR 120
Db 61 BETDVHGKHQSGKWKQMEKPYAFTVHCVRARRHRWKWAQVTFWCPBEQCHLWLQTLR 120

Qy 121 EMLEKLTSPKHLVFINPFGGKGQKRIYERKVAFLPTLASITTDIIIVTEHANOAKETL 180
Db 121 EMLEKLTSPKHLVFINPFGGKGQKRIYERKVAFLPTLASITTDIIIVTEHANOAKETL 180

Qy 181 YEINIDKYDGVICVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPAGS 240
Db 181 YEINIDKYDGVICVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPAGS 240

Qy 241 TDCVCYSTVGTSDAETSALHIVVGDLSLMDVSSVHHNSTLLRYSVSLGFGYFGDIIDKDS 300
Db 241 TDCVCYSTVGTSDAETSALHIVVGDLSLMDVSSVHHNSTLLRYSVSLGFGYFGDIIDKDS 300

Qy 301 EKKRWGLIARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSFDRKPCRCAGCFVCRQSKQ 360
Db 301 EKKRWGLIARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSFDRKPCRCAGCFVCRQSKQ 360

Qy 361 LEEQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
Db 361 LEEQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420

Qy 421 ILIRKCSRFNLRFLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLEKGGKKRF 480
Db 421 ILIRKCSRFNLRFLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLEKGGKKRF 480

Qy 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDHS 537
Db 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDHS 537

RESULT 4
US-10-618-941-121
; Sequence 121, Application US/10618941
; Publication No. US20040197792A1
; GENERAL INFORMATION:
; APPLICANT: MANNING, GERARD
; APPLICANT: CAENEPEEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618,941
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 143
```

OK

```
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 121
; LENGTH: 537
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-618-941-121

Query Match      100.0%; Score 2888; DB 16; Length 537;
Best Local Similarity 100.0%; Pred. No. 3e-278;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGATGAAPLQSLVWVKQRCVSLPARALLRWRSFPGAGAPGADACSVPVSEIIAV 60
Db 1 MGATGAAPLQSLVWVKQRCVSLPARALLRWRSFPGAGAPGADACSVPVSEIIAV 60

Qy 61 BETDVHGKHQSGKWKQMEKPYAFTVHCVRARRHRWKWAQVTFWCPBEQCHLWLQTLR 120
Db 61 BETDVHGKHQSGKWKQMEKPYAFTVHCVRARRHRWKWAQVTFWCPBEQCHLWLQTLR 120

Qy 121 EMLEKLTSPKHLVFINPFGGKGQKRIYERKVAFLPTLASITTDIIIVTEHANOAKETL 180
Db 121 EMLEKLTSPKHLVFINPFGGKGQKRIYERKVAFLPTLASITTDIIIVTEHANOAKETL 180

Qy 181 YEINIDKYDGVICVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPAGS 240
Db 181 YEINIDKYDGVICVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPAGS 240

Qy 241 TDCVCYSTVGTSDAETSALHIVVGDLSLMDVSSVHHNSTLLRYSVSLGFGYFGDIIDKDS 300
Db 241 TDCVCYSTVGTSDAETSALHIVVGDLSLMDVSSVHHNSTLLRYSVSLGFGYFGDIIDKDS 300

Qy 301 EKKRWGLIARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSFDRKPCRCAGCFVCRQSKQ 360
Db 301 EKKRWGLIARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSFDRKPCRCAGCFVCRQSKQ 360

Qy 361 LEEQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
Db 361 LEEQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420

Qy 421 ILIRKCSRFNLRFLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLEKGGKKRF 480
Db 421 ILIRKCSRFNLRFLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDLEKGGKKRF 480

Qy 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDHS 537
Db 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDHS 537

RESULT 5
US-09-969-896-11
; Sequence 11, Application US/09969896
; Publication No. US20030125533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 562
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-969-896-11

Query Match      100.0%; Score 2888; DB 10; Length 562;
Best Local Similarity 100.0%; Pred. No. 3.2e-278;
```

OK



Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGATGAAPLQSVLWVKQORCAVSLPARALLRRWRSFGGAGAGADACSVPSIIIV 60  
Db 26 MGATGAAPLQSVLWVKQORCAVSLPARALLRRWRSFGGAGAGADACSVPSIIIV 85  
Qy 61 EETDVHGKHGSGKQKWKMEKPYAFTVHCVKRARRHRWKWAQVTFWCPEEQCHLWLTQLR 120  
Db 86 EETDVHGKHGSGKQKWKMEKPYAFTVHCVKRARRHRWKWAQVTFWCPEEQCHLWLTQLR 145  
Qy 121 EMLEKLTSPKHLVFINPFGGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180  
Db 146 EMLEKLTSPKHLVFINPFGGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 205  
Qy 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 240  
Db 206 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 265  
Qy 241 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS 300  
Db 266 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS 325  
Qy 301 EKKRWGLARYDFSGKLTFLSHHCYEGTVSFPAQHTVGSPRDRKPCRAGCFVCROSKQ 360  
Db 326 EKKRWGLARYDFSGKLTFLSHHCYEGTVSFPAQHTVGSPRDRKPCRAGCFVCROSKQ 385  
Qy 361 LEEBOKKALYGLEAAEDVEEWOVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420  
Db 386 LEEBOKKALYGLEAAEDVEEWOVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 445  
Qy 421 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDDLKEGGKKRF 480  
Db 446 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDDLKEGGKKRF 505  
Qy 481 GHICSSHPSCCTVSNSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 537  
Db 506 GHICSSHPSCCTVSNSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 562

RESULT 6  
US-10-631-958-11  
; Sequence 11, Application US/10631958  
; Publication No. US20040192580A1  
; GENERAL INFORMATION:  
; APPLICANT: Kossida, Sophia  
; TITLE OF INVENTION: Regulation of human Sphingosine  
; TITLE OF INVENTION: Kinase-Like Protein  
; FILE REFERENCE: 004974.00594  
; CURRENT APPLICATION NUMBER: US/10/631,958  
; PRIOR FILING DATE: 2003-08-01  
; PRIOR APPLICATION NUMBER: US/09/969,896  
; PRIOR FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: US 60/238,005  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: US 60/314,113  
; PRIOR FILING DATE: 2001-08-23  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 562  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-631-958-11

Query Match 100.0%; Score 2888; DB 16; Length 562;  
Best Local Similarity 100.0%; Pred. No. 3.2e-278;  
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGATGAAPLQSVLWVKQORCAVSLPARALLRRWRSFGGAGAGADACSVPSIIIV 60  
Db 26 MGATGAAPLQSVLWVKQORCAVSLPARALLRRWRSFGGAGAGADACSVPSIIIV 85  
Qy 61 EETDVHGKHGSGKQKWKMEKPYAFTVHCVKRARRHRWKWAQVTFWCPEEQCHLWLTQLR 120

Db 86 EETDVHGKHGSGKQKWKMEKPYAFTVHCVKRARRHRWKWAQVTFWCPEEQCHLWLTQLR 145  
Qy 121 EMLEKLTSPKHLVFINPFGGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180  
Db 146 EMLEKLTSPKHLVFINPFGGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 205  
Qy 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 240  
Db 206 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 265  
Qy 241 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS 300  
Db 266 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS 325  
Qy 301 EKKRWGLARYDFSGKLTFLSHHCYEGTVSFPAQHTVGSPRDRKPCRAGCFVCROSKQ 360  
Db 326 EKKRWGLARYDFSGKLTFLSHHCYEGTVSFPAQHTVGSPRDRKPCRAGCFVCROSKQ 385  
Qy 361 LEEBOKKALYGLEAAEDVEEWOVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420  
Db 386 LEEBOKKALYGLEAAEDVEEWOVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 445  
Qy 421 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDDLKEGGKKRF 480  
Db 446 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRVKFKQFTSKHMEDESDDLKEGGKKRF 505  
Qy 481 GHICSSHPSCCTVSNSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 537  
Db 506 GHICSSHPSCCTVSNSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 562

RESULT 7  
US-10-315-597A-2  
; Sequence 2, Application US/10315597A  
; Publication No. US20030162206A1  
; GENERAL INFORMATION:  
; APPLICANT: Sugiura, Masaaki  
; APPLICANT: Kono, Keita  
; APPLICANT: Kohama, Takafumi  
; TITLE OF INVENTION: Ceramide Kinase and DNA Encoding It  
; FILE REFERENCE: 02658CIP/HG  
; CURRENT APPLICATION NUMBER: US/10/315,597A  
; CURRENT FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: JP 2000-178039  
; PRIOR FILING DATE: 2000-06-14  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 537  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-315-597A-2

Query Match 99.7%; Score 2880; DB 14; Length 537;  
Best Local Similarity 99.6%; Pred. No. 1.9e-277;  
Matches 535; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGATGAAPLQSVLWVKQORCAVSLPARALLRRWRSFGGAGAGADACSVPSIIIV 60  
Db 1 MGATGAAPLQSVLWVKQORCAVSLPARALLRRWRSFGGAGAGADACSVPSIIIV 60  
Qy 61 EETDVHGKHGSGKQKWKMEKPYAFTVHCVKRARRHRWKWAQVTFWCPEEQCHLWLTQLR 120  
Db 61 EETDVHGKHGSGKQKWKMEKPYAFTVHCVKRARRHRWKWAQVTFWCPEEQCHLWLTQLR 120  
Qy 121 EMLEKLTSPKHLVFINPFGGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180  
Db 121 EMLEKLTSPKHLVFINPFGGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180  
Qy 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 240  
Db 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 240

```
QY 241 TDCVYSTVGTSDAETSAHIIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYGYGDIIDKS 300
DB |||||
DB 241 TDCVYSTVGTSDAETSAHIIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYGYGDIIDKS 300
QY 301 EKRWGLGARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQ 360
DB |||||
DB 301 EKRWGLGARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQ 360
QY 361 LEEQKALYGLAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 420
DB |||||
DB 361 LEEQKALYGLAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 420
QY 421 ILIRKCSFNFLRFLIRHTNQDQDFTFVEVYRVKXQFTSKHMEDESDLKEGGKRF 480
DB |||||
DB 421 ILIRKCSFNFLRFLIRHTNQDQDFTFVEVYRVKXQFTSKHMEDESDLKEGGKRF 480
QY 481 GHICSHSPCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 537
DB |||||
DB 481 GHICSHSPCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 537

RESULT 8
US-09-784-810A-6
; Sequence 6, Application US/09784810A
; Patent No. US20020082203A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/09/784,810A
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-784-810A-6

Query Match 85.1%; Score 2456.5; DB 9; Length 471;
Best Local Similarity 97.5%; Pred. No. 2.4e-235;
Matches 459; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

QY 78 MEKPYAFTVHCVKRARRHWKWAQVTFWCPBEQLCHLWLQTLREMLEKLTSPKHLVFI 137
DB 1 MEKPYAFTVHCVKRARRHWKWAQVTFWCPBEQLCHLWLQTLREMLEKLTSPKHLVFI 60
QY 138 NPFCKGKGKRIYERKVAPLFTLASITTDII-----VTEHANOAKETLYEINID 186
DB 61 NPFCKGKGKRIYERKVAPLFTLASITTDIIIGNKFYVNYVEVITEHANOAKETLYEINID 120
QY 187 KYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGSTDVCY 246
DB 121 KYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGSTDVCY 180
QY 247 STVGTSDAETSAHIIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYGYGDIIDKS 306
DB 181 STVGTSDAETSAHIIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYGYGDIIDKS 240
QY 307 GLARYDFSLKTLFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQLEEBQ 366
DB 241 GLARYDFSLKTLFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQLEEBQ 300
QY 367 KALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDLIRK 426
DB 301 KALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDLIRK 360
QY 427 STVGTSDAETSAHIIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYGYGDIIDKS 486
DB 181 STVGTSDAETSAHIIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYGYGDIIDKS 420
QY 307 GLARYDFSLKTLFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQLEEBQ 366
DB 241 GLARYDFSLKTLFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQLEEBQ 300
QY 367 KALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDLIRK 426
DB 301 KALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDLIRK 360
```

```
QY 427 SRNFLRFLIRHTNQDQDFTFVEVYRVKXQFTSKHMEDESDLKEGGKRF 486
DB |||||
DB 361 SRNFLRFLIRHTNQDQDFTFVEVYRVKXQFTSKHMEDESDLKEGGKRF 420
QY 487 HPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 537
DB |||||
DB 421 HPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 471

RESULT 9
US-10-876-281-6
; Sequence 6, Application US/10876281
; Publication No. US20050123942A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/10/876,281
; CURRENT FILING DATE: 2004-06-24
; PRIOR FILING DATE: US/09/784,810
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-876-281-6

Query Match 85.1%; Score 2456.5; DB 18; Length 471;
Best Local Similarity 97.5%; Pred. No. 2.4e-235;
Matches 459; Conservative 1; Mismatches 0; Indels 11; Gaps 1;

QY 78 MEKPYAFTVHCVKRARRHWKWAQVTFWCPBEQLCHLWLQTLREMLEKLTSPKHLVFI 137
DB 1 MEKPYAFTVHCVKRARRHWKWAQVTFWCPBEQLCHLWLQTLREMLEKLTSPKHLVFI 60
QY 138 NPFCKGKGKRIYERKVAPLFTLASITTDII-----VTEHANOAKETLYEINID 186
DB 61 NPFCKGKGKRIYERKVAPLFTLASITTDIIIGNKFYVNYVEVITEHANOAKETLYEINID 120
QY 187 KYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGSTDVCY 246
DB 121 KYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGSTDVCY 180
QY 247 STVGTSDAETSAHIIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYGYGDIIDKS 306
DB 181 STVGTSDAETSAHIIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYGYGDIIDKS 240
QY 307 GLARYDFSLKTLFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQLEEBQ 366
DB 241 GLARYDFSLKTLFSLHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQLEEBQ 300
QY 367 KALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDLIRK 426
DB 301 KALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDLIRK 360
QY 427 SRNFLRFLIRHTNQDQDFTFVEVYRVKXQFTSKHMEDESDLKEGGKRF 486
DB 361 SRNFLRFLIRHTNQDQDFTFVEVYRVKXQFTSKHMEDESDLKEGGKRF 420
QY 487 HPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 537
DB |||||
DB 421 HPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 471

RESULT 10
US-09-784-810A-11
```

; Sequence 11, Application US/09784810A  
; Patent No. US20020082203A1  
; GENERAL INFORMATION:  
; APPLICANT: RASTELLI, LUCA  
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING  
; FILE REFERENCE: 10716-08  
; CURRENT APPLICATION NUMBER: US/09/784,810A  
; CURRENT FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: 60/182,360  
; PRIOR FILING DATE: 2000-02-14  
; PRIOR APPLICATION NUMBER: 60/191,261  
; PRIOR FILING DATE: 2000-03-22  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 326  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-784-810A-11

Query Match 56.8%; Score 1640.5; DB 9; Length 326;  
Best Local Similarity 96.3%; Pred. No. 2.9e-154;  
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;  
  
Qy 130 PKHLLVFNPFGKGQKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 178  
Db 1 PKHLLVFNPFGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVNVYVEVITEHANOAKE 60  
  
Qy 179 TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPA 238  
Db 61 TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPA 120  
  
Qy 239 GSTDCVCYSTVGTSDAETSALHIVVGDSLAMDVSSVHHNSTLLRYSVSLGFGFYGDIK 298  
Db 121 GSTDCVCYSTVGTSDAETSALHIVVGDSLAMDVSSVHHNSTLLRYSVSLGFGFYGDIK 180  
  
Qy 299 DSEKKRWLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCRQSK 358  
Db 181 DSEKKRWLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCRQSK 240  
  
Qy 359 QOLEEEQKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 418  
Db 241 QOLEEEQKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300  
  
Qy 419 DLILIRKSRNFLRLIRHTNQDQ 444  
Db 301 DLILIRKSRNFLRLIRHTNQDQ 326

RESULT 11  
US-09-969-896-2  
; Sequence 2, Application US/09969896  
; Publication No. US20030125533A1  
; GENERAL INFORMATION:  
; APPLICANT: Kossida, Sophia  
; TITLE OF INVENTION: Regulation of human Sphingosine  
; FILE REFERENCE: 004974.00594  
; CURRENT APPLICATION NUMBER: US/09/969,896  
; CURRENT FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: US 60/238,005  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: US 60/314,113  
; PRIOR FILING DATE: 2001-08-23  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 326  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-969-896-2

Query Match 56.8%; Score 1640.5; DB 10; Length 326;  
Best Local Similarity 96.3%; Pred. No. 2.9e-154;  
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;  
  
Qy 130 PKHLLVFNPFGKGQKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 178  
Db 1 PKHLLVFNPFGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVNVYVEVITEHANOAKE 60  
  
Qy 179 TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPA 238  
Db 61 TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPA 120  
  
Qy 239 GSTDCVCYSTVGTSDAETSALHIVVGDSLAMDVSSVHHNSTLLRYSVSLGFGFYGDIK 298  
Db 121 GSTDCVCYSTVGTSDAETSALHIVVGDSLAMDVSSVHHNSTLLRYSVSLGFGFYGDIK 180  
  
Qy 299 DSEKKRWLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCRQSK 358  
Db 181 DSEKKRWLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCRQSK 240  
  
Qy 359 QOLEEEQKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 418  
Db 241 QOLEEEQKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300  
  
Qy 419 DLILIRKSRNFLRLIRHTNQDQ 444  
Db 301 DLILIRKSRNFLRLIRHTNQDQ 326

RESULT 12  
US-10-631-958-2  
; Sequence 2, Application US/10631958  
; Publication No. US20040192580A1  
; GENERAL INFORMATION:  
; APPLICANT: Kossida, Sophia  
; TITLE OF INVENTION: Regulation of human Sphingosine  
; FILE REFERENCE: 004974.00594  
; CURRENT APPLICATION NUMBER: US/10/631,958  
; CURRENT FILING DATE: 2003-08-01  
; PRIOR APPLICATION NUMBER: US/09/969,896  
; PRIOR FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: US 60/238,005  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: US 60/314,113  
; PRIOR FILING DATE: 2001-08-23  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 326  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-631-958-2

Query Match 56.8%; Score 1640.5; DB 16; Length 326;  
Best Local Similarity 96.3%; Pred. No. 2.9e-154;  
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;  
  
Qy 130 PKHLLVFNPFGKGQKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 178  
Db 1 PKHLLVFNPFGKGQKRIYERKVAFLFTLASITTDIIIGNKFYVNVYVEVITEHANOAKE 60  
  
Qy 179 TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPA 238  
Db 61 TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLPVSSLRIGIIPA 120  
  
Qy 239 GSTDCVCYSTVGTSDAETSALHIVVGDSLAMDVSSVHHNSTLLRYSVSLGFGFYGDIK 298  
Db 121 GSTDCVCYSTVGTSDAETSALHIVVGDSLAMDVSSVHHNSTLLRYSVSLGFGFYGDIK 180  
  
Qy 299 DSEKKRWLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCRQSK 358  
Db 181 DSEKKRWLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCRQSK 240

QY 359 QOLEBEQKALYGLAEAEVVEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSS 418  
DB 241 QOLEBEQKALYGLAEAEVVEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300  
QY 419 DLILIRKCSRNFNRLFLIRHTNQDQ 444  
DB 301 DLILIRKCSRNFNRLFLIRHTNQDQ 326

## RESULT 13

US-10-876-281-11  
; Sequence 11, Application US/10876281  
; Publication No. US20050123942A1  
; GENERAL INFORMATION:  
; APPLICANT: RASTELLI, LUCA  
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING  
; FILE REFERENCE: 10716-08  
; CURRENT APPLICATION NUMBER: US/10/876,281  
; CURRENT FILING DATE: 2004-06-24  
; PRIOR APPLICATION NUMBER: US/09/784,810  
; PRIOR FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: 60/182,360  
; PRIOR FILING DATE: 2000-02-14  
; PRIOR APPLICATION NUMBER: 60/191,261  
; PRIOR FILING DATE: 2000-03-22  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 326  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-876-281-11

Query Match 56.8%; Score 1640.5; DB 18; Length 326;  
Best Local Similarity 96.3%; Pred. No. 2.9e-154;  
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;  
QY 130 PKHLLVFVFNPFQGGQKRIYERKVAFLFTLASITTDII-----VTEHANOAKE 178  
DB 1 PKHLLVFVFNPFQGGQKRIYERKVAFLFTLASITTDIIIGNKFVYVVEVITEHANOAKE 60  
QY 179 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPA 238  
DB 61 TLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPSSLRIGIIPA 120  
QY 239 GSTDCVCYSTVGTSDAETSALHIVGDSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIK 298  
DB 121 GSTDCVCYSTVGTSDAETSALHIVGDSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIK 180  
QY 299 DSEKKRWLGARYDFSGKTLFSLHHCYEGTGVSLFPAQHTVGSFPRDKPCRAGCFVCROSK 358  
DB 181 DSEKKRWLGARYDFSGKTLFSLHHCYEGTGVSLFPAQHTVGSFPRDKPCRAGCFVCROSK 240  
QY 359 QOLEBEQKALYGLAEAEVVEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSS 418  
DB 241 QOLEBEQKALYGLAEAEVVEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300  
QY 419 DLILIRKCSRNFNRLFLIRHTNQDQ 444  
DB 301 DLILIRKCSRNFNRLFLIRHTNQDQ 326

## RESULT 14

US-09-784-810A-29  
; Sequence 29, Application US/09784810A  
; Patent No. US20020082203A1  
; GENERAL INFORMATION:  
; APPLICANT: RASTELLI, LUCA  
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING  
; FILE REFERENCE: 10716-08

; CURRENT APPLICATION NUMBER: US/09/784,810A  
; CURRENT FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: 60/182,360  
; PRIOR FILING DATE: 2000-02-14  
; PRIOR APPLICATION NUMBER: 60/191,261  
; PRIOR FILING DATE: 2000-03-22  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 29  
; LENGTH: 182  
; TYPE: PRT  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: 80432911  
US-09-784-810A-29

Query Match 34.1%; Score 986; DB 9; Length 182;  
Best Local Similarity 100.0%; Pred. No. 2e-89;  
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 42 AGAFGADACSVPVSEIIIAVEETDVHGKHQSGKQKMEKPYAFTVHCVKRARRHRWKWAQ 101  
DB 1 AGAFGADACSVPVSEIIIAVEETDVHGKHQSGKQKMEKPYAFTVHCVKRARRHRWKWAQ 60  
QY 102 VTFWCPBEQLCHLWLTREMLEKLTSPKHLVFINPFGGKGQKRIYERKVAFLFTLA 161  
DB 61 VTFWCPBEQLCHLWLTREMLEKLTSPKHLVFINPFGGKGQKRIYERKVAFLFTLA 120  
QY 162 SITTDIIIVTEHANOAKETLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNH 221  
DB 121 SITTDIIIVTEHANOAKETLYEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNH 180  
QY 222 PR 223  
DB 181 PR 182

## RESULT 15

US-10-876-281-29  
; Sequence 29, Application US/10876281  
; Publication No. US20050123942A1  
; GENERAL INFORMATION:  
; APPLICANT: RASTELLI, LUCA  
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING  
; FILE REFERENCE: 10716-08  
; CURRENT APPLICATION NUMBER: US/10/876,281  
; CURRENT FILING DATE: 2004-06-24  
; PRIOR APPLICATION NUMBER: US/09/784,810  
; PRIOR FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: 60/182,360  
; PRIOR FILING DATE: 2000-02-14  
; PRIOR APPLICATION NUMBER: 60/191,261  
; PRIOR FILING DATE: 2000-03-22  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 29  
; LENGTH: 182  
; TYPE: PRT  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: 80432911  
US-10-876-281-29

Query Match 34.1%; Score 986; DB 18; Length 182;  
Best Local Similarity 100.0%; Pred. No. 2e-89;  
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 42 AGAFGADACSVPVSEIIIAVEETDVHGKHQSGKQKMEKPYAFTVHCVKRARRHRWKWAQ 101  
DB 1 AGAFGADACSVPVSEIIIAVEETDVHGKHQSGKQKMEKPYAFTVHCVKRARRHRWKWAQ 60  
QY 102 VTFWCPBEQLCHLWLTREMLEKLTSPKHLVFINPFGGKGQKRIYERKVAFLFTLA 161

Db	61	VTFWCPPEQLCHLWLQTLREMLEKLTSRPKHLLVFINPFGGKGQKRIYERKVAFLFLA	120
Qy	162	SITTDIIIVTEHANQAKETLYEINIDKYDGIVCVGDCGMFSEVLHGLIGRTORSAGVDQNH	221
Db	121	SITTDIIIVTEHANQAKETLYEINIDKYDGIVCVGDCGMFSEVLHGLIGRTORSAGVDQNH	180
Qy	222	PR 223	
Db	181	PR 182	

Search completed: September 3, 2005, 04:32:35  
Job time : 115.269 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 04:25:44 ; Search time 99.8632 Seconds  
(without alignments)  
2753.633 Million cell updates/sec

Title: US-10-631-958-10  
Perfect score: 2888  
Sequence: 1 NGATGAEPLOSLVWVKQOR.....QLVRLFARGIEENPKDPDHS 537

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2888	100.0	537	1 CEK1_HUMAN	Q8tct0 homo sapien
2	2428	84.1	531	1 CEK1_MOUSE	Q8k4q7 mus musculus
3	1904	65.9	409	2 Q62PK5	Q6zpk5 mus musculus
4	1811	62.7	339	2 Q6NX59	Q6nx59 homo sapien
5	1713	59.3	572	2 Q6GLV1	Q6glv1 xenopus lae
6	1495	51.8	485	2 Q6GMF3	Q6gmf3 xenopus lae
7	801.5	27.8	687	2 Q9VNA6	Q9vna6 drosophila
8	763	26.4	487	2 Q95S15	Q95s15 drosophila
9	666.5	23.1	410	2 Q7PRA8	Q7pra8 anopheles g
10	648	22.4	608	2 Q6USK2	Q6usk2 arabidopsis
11	619	21.4	700	2 Q6H6H1	Q6h6h1 oryza sativ
12	583.5	20.2	533	2 Q9LU45	Q9lu45 arabidopsis
13	576.5	20.0	532	2 Q6UZF6	Q6ufz6 homo sapien
14	567.5	19.7	549	2 Q9TZ11	Q9tzi1 caenorhabdi
15	480	16.6	586	2 Q949C3	Q949c3 oryza sativ
16	428	14.8	485	2 Q817L1	Q817l1 arabidopsis
17	417.5	14.5	1240	2 Q65419	Q65419 arabidopsis
18	369.5	12.8	763	2 Q9LRB0	Q9lrb0 arabidopsis
19	368.5	12.8	480	2 Q8H350	Q8h350 oryza sativ
20	349	12.1	732	2 Q9PHL3	Q9phl3 arabidopsis
21	343	11.9	280	2 Q6ZP59	Q6zp59 homo sapien
22	331.5	11.5	446	2 Q84S01	Q84s01 oryza sativ
23	330.5	11.4	757	2 Q84HY9	Q84hy9 oryza sativ
24	330.5	11.4	757	2 Q7KCS9	Q7kcs9 oryza sativ
25	323	11.2	685	2 Q7Q1P4	Q7qlp4 anopheles g
26	314.5	10.9	748	2 Q7XN57	Q7xn57 oryza sativ
27	310.5	10.8	641	2 Q9VYH8	Q9vyh8 drosophila
28	298.5	10.3	661	2 Q9VZW0	Q9vzw0 drosophila
29	291.5	10.1	458	2 Q14159	Q14159 schizosacch
30	271.5	9.4	654	1 SPH2_HUMAN	Q6ayb2 homo sapien
31	267.5	9.3	616	2 Q6AYB2	Q6ayb2 rattus norv

32	267.5	9.3	617	1 SPH2_MOUSE	Q9jia7 mus musculus
33	264.5	9.2	384	1 SPH1_HUMAN	Q9nyal homo sapien
34	264.5	9.2	384	2 Q96HV8	Q96hv8 homo sapien
35	264.5	9.2	398	2 Q96GK1	Q96gk1 homo sapien
36	264.5	9.2	470	2 Q8N632	Q8n632 homo sapien
37	260	9.0	388	2 Q88886	Q88886 mus musculus
38	257.5	8.9	624	2 Q86KP9	Q86kf9 dictyosteli
39	254	8.8	381	2 Q91ZN3	Q91zn3 mus musculus
40	254	8.8	382	2 Q8C115	Q8c115 mus musculus
41	254	8.8	504	2 Q88885	Q88885 mus musculus
42	253.5	8.8	579	2 Q74ZE3	Q74ze3 ashbya goss
43	248.5	8.6	687	2 Q06147	Q06147 saccharomyc
44	245.5	8.5	383	2 Q642F6	Q642f6 rattus norv
45	245.5	8.5	423	2 Q7JM91	Q7jm91 caenorhabdi

#### ALIGNMENTS

RESULT 1

ID	CEK1_HUMAN	STANDARD;	PRT;	537 AA.
AC	Q8TCT0; Q9BYB3; Q9UGES;			
DT	10-OCT-2003 (Rel. 42, Created)			
DT	10-OCT-2003 (Rel. 42, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Ceramide kinase (EC 2.7.1.138) (Acylsphingosine kinase) (hCERK) (Lipid			
DE	kinase 4) (LK4).			
GN	Name=CERK; Synonyms=KIAA1646;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.			
RC	TISSUE=Leukemia;			
RX	MEDLINE=22075121; PubMed=11956206; DOI=10.1074/jbc.M201535200;			
RA	Suglura M., Kono K., Liu H., Shimizugawa T., Minekura H., Spiegel S.,			
RA	Kohama T.;			
RT	"Ceramide kinase, a novel lipid kinase. Molecular cloning and			
RT	functional characterization."			
RL	J. Biol. Chem. 277:23294-23300(2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Van Veldhoven P.P.;			
RT	"A search for lipid kinases."			
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=20057165; PubMed=10591208; DOI=10.1038/990031;			
RA	Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,			
RA	Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,			
RA	Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,			
RA	Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,			
RA	Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,			
RA	Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,			
RA	Conroy D., Corby N.R., Coville S.J., Cox A.V., Davis J., Dawson E.,			
RA	Dhami P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,			
RA	Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,			
RA	Gilbert J.G.R., Goward M.E., Graham D.V., Griffiths M.N.D., Hall C.,			
RA	Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,			
RA	Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,			
RA	Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,			
RA	Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., McCann O.T.,			
RA	Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,			
RA	Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,			
RA	Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,			
RA	Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,			
RA	Soderlund C., Spraggon L., Steward C.A., Sulston J.E., Swann R.M.,			
RA	Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,			
RA	Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,			
RA	Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,			
RA	Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,			

RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuayama S.,  
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,  
RA Dorman A., Fang F., Fu Y., Hua A., Kenton S., Lai H., Lao H.I.,  
RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,  
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,  
RA Wang C., Wang Z., White J., Willingham D., Wu H., Yao Z.,  
RA Zhan M., Zhang G., Chisoso S., Murray J., Miller N., Minx P.,  
RA Fulton R., Johnson D., Benis G., Bentley D., Bradshaw H., Bourne S.,  
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,  
RA Hinds K., Kemp K., Latreille P., Layman D., Ozeresky P., Rohlfing T.,  
RA Scheet P., Walker C., Wamsley A., Wohldmann P., Pepin K., Nelson J.,  
RA Korf I., Bedell J.A., Hillier L.M., Mardis E., Waterston R.,  
RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,  
RA Budarf M.L., McDermaid H.E., Johnson A., Wong A.C.C., Morrow B.E.,  
RA Edelmann L., Kim U.J., Shizuwa H., Simon M.I., Dumanek J.P.,  
RA Peyrard M., Kedra D., Seroussi E., Fransson I., Tapia I., Bruder C.E.,  
RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,  
RA Khan A.S., Lane L., Tiliahun Y., Wright H.;  
RA "The DNA sequence of human chromosome 22.";  
RT Nature 402:489-495(1999).  
RN [4]  
RP SEQUENCE OF 57-537 FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=21156230; PubMed=11258795;  
RA Hirosewa M., Nagase T., Murahashi Y., Kikuno R., Ohara O.;  
RA "Identification of novel transcribed sequences on human chromosome 22  
RT by expressed sequence tag mapping.";  
RL DNA Res. 8:1-9(2001).  
CC -1- FUNCTION: Catalyzes specifically the phosphorylation of ceramide  
CC to form ceramide 1-phosphate. Acts efficiently on natural and  
CC analog ceramides (C6, C8, C16 ceramides, and C8-dihydroceramide),  
CC to a lesser extent on C2-ceramide and C6-dihydroceramide, but not  
CC on other lipids, such as various sphingolipids.  
CC -1- CATALYTIC ACTIVITY: ATP + ceramide = ADP + ceramide 1-phosphate.  
CC -1- COFACTOR: Calcium and magnesium.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.  
CC -1- TISSUE SPECIFICITY: High level expression in heart, brain,  
CC skeletal muscle, kidney and liver; moderate in peripheral blood  
CC leukocytes and thymus; very low in spleen, small intestine,  
CC placenta and lung.  
CC -1- MISCELLANEOUS: Optimal pH is 6.0-7.5.  
CC -1- SIMILARITY: Contains 1 DAGKc domain.  
CC -1- CAUTION: Ref.3 sequence differs from that shown due to erroneous  
CC gene model prediction. An additional exon may exist between amino  
CC acid positions 168 and 169.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AB079066; BAC01154.1; -;  
DR EMBL; AJ457828; CAD29884.1; -;  
DR EMBL; AL096766; CAB62977.1; ALT SEQ.  
DR EMBL; AL118516; -; NOT ANNOTATED\_CDS.  
DR EMBL; AB051433; BAB33316.1; -;  
DR Genew; HGNC:19256; CERK.  
DR GO; GO:0000299; C: integral to membrane of membrane fraction; IDA.  
DR GO; GO:0004685; F: calcium/calmodulin-dependent protein kinase. .; IDA.  
DR GO; GO:0001729; F: ceramide kinase activity; IDA.  
DR GO; GO:0000287; F: magnesium ion binding; IDA.  
DR GO; GO:0006672; F: ceramide metabolism; TAS.  
DR InterPro; IPR001206; DAGKc.  
DR InterPro; IPR001849; PH.  
DR InterPro; IPR011036; PH related.  
DR Pfam; PF00781; DAGK\_cat; 1.  
DR ProDom; PD005043; DAGKc; 1.  
DR SMART; SM00046; DAGKc; 1.  
DR SMART; SM00233; PH; 1.  
KW Calcium; Kinase; Magnesium; Transferase.

FT DOMAIN 132 278 DAGKc.  
SQ SEQUENCE 537 AA; 59977 MW; 3DBFC0ED8D679F7F CRC64;  
Query Match 100.0%; Score 2888; DB 1; Length 537;  
Best Local Similarity 100.0%; Pred. No. 5e-233;  
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MGATGAAPLQSLVWVKQKQCAVSLPARALLRWSPGFGAGAGADACSVPVSEIIAV 60  
Db 1 MGATGAAPLQSLVWVKQKQCAVSLPARALLRWSPGFGAGAGADACSVPVSEIIAV 60  
Qy 61 EETDVHGKHGSGKQKQKMEKPYAFTVHCIVRARHRHKKWAQVTFWCPBEOLCHLWLOTLR 120  
Db 61 EETDVHGKHGSGKQKQKMEKPYAFTVHCIVRARHRHKKWAQVTFWCPBEOLCHLWLOTLR 120  
Qy 121 EMLEKLTSTRPKHLVFINPFGKGQKRIYERKVAFLFTLASITTTDIIIVTEHANAOKETL 180  
Db 121 EMLEKLTSTRPKHLVFINPFGKGQKRIYERKVAFLFTLASITTTDIIIVTEHANAOKETL 180  
Qy 181 YEINIDKYDGIKVGCGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 240  
Db 181 YEINIDKYDGIKVGCGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 240  
Qy 241 TDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLLGTFGFGDIIDKS 300  
Db 241 TDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLLGTFGFGDIIDKS 300  
Qy 301 EKKRWGLARYDFSLKTLFLSHHCYEGTVSFLPAQHTVSGSPDRKPCRCACGFCVCRQSKQ 360  
Db 301 EKKRWGLARYDFSLKTLFLSHHCYEGTVSFLPAQHTVSGSPDRKPCRCACGFCVCRQSKQ 360  
Qy 361 LEEBQKALYGLEAAEDVEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420  
Db 361 LEEBQKALYGLEAAEDVEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420  
Qy 421 ILIRKCSRNFPLRLHNTNQOQDFTFVEVVRVKKQFTSKHMEDESDLKEGGKKRF 480  
Db 421 ILIRKCSRNFPLRLHNTNQOQDFTFVEVVRVKKQFTSKHMEDESDLKEGGKKRF 480  
Qy 481 GHICSSHPSCCCTVSNSSNCDGEVLHSPAIEVRVHCOLVRLFARGLEENPKPDSSH 537  
Db 481 GHICSSHPSCCCTVSNSSNCDGEVLHSPAIEVRVHCOLVRLFARGLEENPKPDSSH 537  
RESULT 2  
CEK1\_MOUSE STANDARD; PRT; 531 AA.  
AC Q8K4Q7;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Ceramide kinase (EC 2.7.1.138) (Acylsphingosine kinase) (mCERK).  
GN Name=Cerk;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=22075121; PubMed=11956206; DOI=10.1074/jbc.M201535200;  
RA Sugiura M., Kono K., Liu H., Shimizuawa T., Minekura H., Spiegel S.,  
RA Kohama T.;  
RT "Ceramide kinase, a novel lipid kinase. Molecular cloning and  
RT functional characterization.";  
RL J. Biol. Chem. 277:23294-23300(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Heart, and Thymus;  
RX MEDLINE=22354683; PubMed=12456851; DOI=10.1038/nature01266;  
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
RA Nikaudo I., Osato N., Saito R., Suzuki H., Yananaka I., Kiyosawa H.,  
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,



QY		1	MGATGAAPLOSVLWVKKOQRCAVSLEPARALLLRWRSPGAGAPGADACSVPVSEIIAV	60
Dd		1	MGAMGAAPLHSLVLWKRRRCVLSLEPARALLLRWRSPGPGAPCADARSVLVSEIIAV	60
QY		61	BETDVHGKHGGSKWKQKXEPYAFVHCVRARRHAKWAQVFWCPBOLCHLWLQTLR	120
Dd		61	EKDKCEKHSGSRWHKKNPFATFHRVKRVRRHHRWKWARVTFWSADBOLCHLWLQTLR	120
QY		121	EMLEKLTSRPKHLVFINPFGKGQGKRIVYERKVAFLPTLASITTDIIVTEHANQAETL	180
Dd		121	GJLESLSRPHLLVFINPFGKGQGKRIVEKVAPLFTLASITTEIITEHANQAETL	180
QY		181	YEINIDKYDGIUCVCGDMFSEVLHGLIGRTQSAGVDQNHPRAVLVPSLSRIGIIPAGS	240
Dd		181	YEINTDSYDGIUCVCGDMFSEVLHGVIGRTQOSAGIDPNHPRAVLVPSTLRIGIIPAGS	240
QY		241	TDCVCYSTVGTSDAETSALMHTVGDLSAMDVSSVHNSTILLRSVSLGLGVFGDI IKDS	300
Dd		241	TDCVCYSTVGNDSETALSMTIHIGDSLDAIDVSSHVTNLTLLRSVSLGLGVFGDI IKDS	300
QY		301	EKKRWLGRLARYDFSGLKTFLSHHCYEGTVSFPLPAQTIVGSPDRPKPCRAGCFVCROSKQ	360
Dd		301	EKKRWMLGRVDFSGLKTFLSHQYEGTSLFPLPAQTIVGSPDRNKPCRAGCFVCROSKQ	360
QY		361	LEEEOKKALYGLEAAEDVEEVQVCGFKLA INATNMSCARRSPRGLSPAHLGDGSSDL	420
Dd		361	LEEEOKKALYGLENAEEVMQVTCGFKLA INATNMSCAPRSPGGLSPAHLGDGSSDL	420
QY		421	ILIRKCSFPNFLRIPIHTNQODQDFTFVVYRVKKFQFTSKHMEDESDLKEGCKKRF	480
Dd		421	ILIRKCSFPNFLRIPIHTNQODQDFTFVVYRVKKFHTSKHVEDENDSKEQBQKXF	480
QY		481	GHCSSHPSCCTVSNSSWNCDEGLVHSPAIERVHVHCOLVRLPARGIEE	529
Dd		481	GKICWDPSCTCSASRSSWNCDEVMSPAIERVHVHCOLVRLPARGIEE	529
RESULT 3				
Q6ZPKS				
ID	Q6ZPKS	PRELIMINARY;	PRT;	409 AA.
AC	Q6ZPKS;			
DT	05-JUL-2004 (TrEMBRel. 27, Created)			
DT	05-JUL-2004 (TrEMBRel. 27, Last annotation update)			
DT	05-JUL-2004 (TrEMBRel. 27, Last annotation update)			
DN	MKIAA1646 protein (Fragment).			
GE	Name=mKIAA1646;			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
NB	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
TX	TISSUE=Embryonic tail;			
RC	PubMed=14621295;			
RA	Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,			
RA	Saga Y., Nagase T., Ohara O., Koga H.;			
RT	"prediction of the coding sequences of mouse homologues of KIAA gene:			
RT	III. the complete nucleotide sequences of 500 mouse KIAA-homologous			
RT	cDNAs identified by screening of terminal sequences of cDNA clones			
RT	randomly sampled from size-fractionated libraries.";			
RL	DNA Res. 10:167-180(2003).			
DR	EMBL; AK129416; BAC98226.1; "			
DR	GO; GO:0004143; F:diacylglycerol kinase activity; IEA.			
DR	GO; GO:0007205; P:protein kinase C activation; IEA.			
DR	InterPro; IPR001206; DAGKC.			
DR	Pfam; PF00781; DAGK_cat; 1.			
DR	ProDom; PD005043; DAGKC; 1.			
DR	SMART; SM00046; DAGKC; 1.			
FT	NON TER			
SEQ	SEQUENCE	409 AA;	45697 MW;	810B815B37E26808 CRC64;
Query Match				
Best Local Similarity				
Query	Match			
Score	1904;			
DB 2;	Length			
409;				
Pred. No.	9.2e-151;			

Matches	352;	Conservative	26;	Mismatches	24;	Indels	0;	Gaps	0;			
Qy	128	SRPKHLLVF	INPFQGGKGGKRI	YERKVA	PLFTLASITTDIT	IVTTEH	ANAKETL	YEINIDK	187			
Db	6	SRPKHLLVF	INPFQGGKGGKRI	YEKTVAP	LFTLASITTEII	ITTEH	ANAKETL	YEINTDS	65			
Qy	188	YDGIVCVGGDGMF	SEVLHGLIGRTOR	SAGVDQNH	PRAVLVPSS	LIRIGI	IIPAGSTDC	VCYS	247			
Db	66	YDGIVCVGGDGMF	SEVLHGVIGRT	QQSAGIDPN	HPRAVLVPST	LIRIGI	IIPAGSTDC	VCYS	125			
Qy	248	TVGTSDAET	SALHTITVCGDS	LAMDYSSVHHN	STLLARYSVSL	LGYGFYGD	IIKDSEKRWLG	307				
Db	126	TVGTNDT	ETSA	LHIIIGDS	LAI	DYSSVHHN	STLLARYSVSL	LGYGFYGD	IIKDSEKRWLG	185		
Qy	308	LARYDFSG	LKTF	LSHHCYEG	TVSFLP	QAQHTVGS	PRDKPC	RAGCFVCR	SQSQOLEE	OKK	367	
Db	186	LVRDYFSG	LKTF	LSHQYEG	TVSFLP	QAQHTVGS	PRDNKPC	RAGCFVCR	SQSQOLEE	BKK	245	
Qy	368	ALYGLEAAE	VEEVQWV	CGKFLA	INATNM	SCACRRSP	GLSPAAH	LDGSSD	LILRKCS	427		
Db	246	ALYGLENAE	VEEVQWV	CGKFLA	INATNM	SCACRRSP	GLSPFAHL	LDGSSD	LILRKCS	305		
Qy	428	RNFPLRL	IIRHTNQ	ODQDF	FTFVE	YVRVKK	FOFTSK	MEDESD	LDLKEG	KRFGH	CSH	487
Db	306	RNFPLRL	IIRHTNQ	ODQDF	FTFVE	YVRVKK	FOFTSK	MEDESD	LDLKEG	KRFGH	CKDR	365
Qy	488	PSCC	TVSNSS	WNC	DGE	VLH	SPAE	IVR	VHCOL	VL	FARGIEE	529
Db	366	PSCTCS	ASR	SSWNC	DGE	VH	SPAE	IVR	VHCOL	VL	FARGIEE	407

RESULT 4

Q6NX59	PRELIMINARY;	PRT;	339 AA.
ID	Q6NX59		
AC	Q6NX59;		
DT	05-JUL-2004 (TrEMBLrel. 27, Created)		
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)		
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)		
DE	CRK protein.		
GN	NamesCRK;		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RI	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Periphereal Nervous System;		
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Scapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,		
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dixon M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,		
RA	Krzywinski M.I., Skalska U., Smalil D.E., Schnerk A., Schein J.E.,		
RA	Jones S.J., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length human		
RT	and mouse cDNA sequences."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Periphereal Nervous System;		
RL	Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases		

DR	EWBL; BC067255; AAU67255.1; -;
DR	GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
DR	GO; GO:0007205; P:protein kinase C activation; IEA.
DR	InterPro; IPR001206; DAGKC.
DR	ProDom; PD005043; DAGKC; 1.
SQ	SEQUENCE   339 AA; 37780 MW; A4C2ACDFF2E6F3D0 CRC64;
 Query Match                 62.7%; Score 1811; DB 2; Length 339; Best Local Similarity      100.0%; Pred. No. 4.e-143; Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	199 MFSEVHLGLIGRTORSAGVDQHRAVLVPSSLRGIIPAGSTDCVCYSTGTSDAETS <span style="float:right;">258</span>
Dd	1 MFSEVHGLIGRTORSAGVDQHRAVLVPSSLRGIIPAGSTDCVCYSTGTSDAETS <span style="float:right;">60</span>
Qy	259 LHIIVVGDSLAMDVS <del>VHHN</del> TLLRYSVSLLCGYGFYGDIIKDSEKKRWLGARYDP <del>SGLKT</del> 318
Dd	61 LHIIVVGDSLAMDVS <del>VHHN</del> TLLRYSVSLLCGYGFYGDIIKDSEKKRWLGARYDP <del>SGLT</del> 120
Qy	319 FLSHHCYEVTGSVFPAQHTVGPSRDRLKPCRCAGCFVCRQS <del>KOOLEBEQKALYGLEAAEDV</del> 378
Dd	121 FLSHHCYEVTGSVFPAQHTVGPSRDRLKPCRCAGCFVCRQS <del>KOOLEBEQKALYGLEAAEDV</del> 180
Qy	379 EEWQVCGKFIAINATNNSCACRS <del>PRLSPAAHLGDGSSDLLILRKCSRNFRLFIRH</del> 438
Dd	181 EEWQVCGKFIAINATNNSCACRS <del>PRLSPAAHLGDGSSDLLILRKCSRNFRLFIRH</del> 240
Qy	439 TNQQDDQDFTFVEVYRVVKKFOFTSKHWDEDSDLKEGGKKRFGHICSSHPSCCCTVSNS <span style="float:right;">S</span>
Dd	241 TNQQDDQDFTFVEVYRVVKKFOFTSKHWDEDSDLKEGGKKRFGHICSSHPSCCCTVSNS <span style="float:right;">300</span>
Qy	499 WNCDEVILHSPIAEVRVCHQLVRLFARGIEENPKPDSHS 537
Dd	301 WNCDEVILHSPIAEVRVCHQLVRLFARGIEENPKPDSHS 339

RESULT 5	Q6GLV1	PRELIMINARY;	PRT;	572 AA.
ID	Q6GLV1			
AC	Q6GLV1			
DC	05-JUL-2004 (TrEMBLrel. 27, Created)			
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)			
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)			
DE	MG84197 protein.			
GN	Name=MG84197;			
OS	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;			
OC	Xenopodinae; Xenopus.			
OX	NCBI_TaxID=8355;			
RX	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Scapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalon D.K., Munzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,			
RA	Krzywinski M.I., Skalka U., Smailus D.E., Schnerch A., Schein J.E.,			
RA	Jones S.J., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences."			

```
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Klein S., Gerhard D.S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC074350; AAH74350.1; -.
DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
DR InterPro; IPR001206; DAGKC.
DR Pfam; PF00781; DAGK_cat; 1.
DR ProDom; PD005043; DAGKC; 1.
DR SMART; SM00046; DAGKC; 1.
SQ SEQUENCE 572 AA; 64899 MW; E3D22DB2C93B5B77 CRC64;

Query Match 59.3%; Score 1713; DB 2; Length 572;
Best Local Similarity 58.4%; Pred. No. 1.4e-134;
Matches 320; Conservative 81; Mismatches 117; Indels 30; Gaps 4;

QY 10 LQSVLWVKQRCQCAVSLPARALLRWRSPPGAG-APG----- 46
DQ 12 LCSVLVSKRQCEVTLDPGRSLLS-WRELPRRCKDRSRPGICMTRVFHQARGLVNGFTH 70
QY 47 ---ADACSVPVSEIIAIVETDVHGKQSGKQKMEKPYAFTHVCVKRARRHRKWAQVT 103
DQ 71 QPFGSGVTVPVTEIVSGEAEIDEKYNSMKWQHICKPRAFTVRVQVRKRWCKEVT 130
QY 104 FWCPEQLCHLWQLTREMLEKLTSPKHLIVFNPFPGKGGKGIYERKVPALFTLASI 163
DQ 131 FWCDEQLFYQWLQAFHDLLEQOQTHRPKNLLVYINPYGKKRGKQIYENKVPALFSAAGI 190
QY 164 TTDIVTTHANQAKETLVEINIDKVDGIVCVGDCGMFSEVLHGLIGRTQRSAGVDONHPR 223
DQ 191 CADVIVTEYANHARDHLIDANLEKTDGVVCGDGMFSEVLHGLIVRMQKDSVDVHNHPS 250
QY 224 AVLVPSSLRIGIIPAGSDCYSTVGTSDAETSALHIWVGDLSAMDVSSVHHNSTLLRY 283
DQ 251 AQLSRCNMRIGIIPAGSDCYATVGTINDPETSALHIILGDCQPLDVCSVHYKRTFLKY 310
QY 284 SVSLGCGFYGDIIKDSKKRWGLIARYDFSLGKTLFSLHCYEGTVSFLPAQHTVGSPPD 343
DQ 311 SVSLGCGFYGDVLGSEKRWLGPARVDVSGFKTLFTHHCYEGSVSPQAKWVLGSPRD 370
QY 344 RKPCRAGCFVCRQSQQLLEEKKALYGLEAAEDVEEQVCGKFLAINATNMSCACPRS 403
DQ 371 QTCTSGGCIYCRQSKQDEQQTQACSEHREQDDMTTIGRFMAINVSMSCACPRS 430
QY 404 PRGLSPAALHGDGSSDLILIRKSRFNLRLIRHTNQDQDFTFVEVYRKKFQFTSK 463
DQ 431 PNGLSPAALHADGSDADLLIVRKSRDLFLRLIRHTSNKQDFDFVEVYRKKFQFTPK 490
QY 464 HMEDED---SDLKEGKKRFGHICSSHPSCCCTVSNSSNCCDGEVLHSPAIEVRVHCOLV 520
DQ 491 HPEDENESTDILGKKNFSQICTDHPSCGNHVNISWNCCDGETLDQTAIEMRVHCQLI 550
QY 521 RLFAEGIE 528
DQ 551 KLFARGIE 558

RESULT 6
QSGMF3
ID Q6GMF3 PRELIMINARY; PRT; 485 AA.
AC Q6GMF3;
```

Qy 164 TTDIIIVTEHANOAKETLYEINIDKYDGVCGDGMFSEVLHGLIGRTORSAGVDQNHPR 223  
Db 191 CADVIVTEYANHARDNLYENLVKYGVCVCGDGFSEVLHGLIGRTORSAGVDHNNPN 250  
Qy 224 AVLVPSSIRIGIIPAGSTDCVYCTVGTSDAETSALHIVGDSLAMDVSSVHHNSTLLRY 283  
Db 251 AHLQSCNIRIGIIPAGSTDCIYATVGINDPETSALHIIIGDQCPDLVSCVHHNRTTFLKY 310  
Qy 284 SVSLGGLGYFYGDIIKDEKRWGLGARYDFSGGLTKFLSHHCYEGTGFPLPAQHTVGSFPRD 343  
Db 311 SVSLGGLGYFYGDVLKTEKRWGLGARYDYVSGCKTFLTHHCYEGSVFQPAKXVWLGSPRD 370  
Qy 344 RKPACAGFCVCRQSQOLEEOKKALYGLE-AAEDVEEMQVQVCGKFLAINATNMSCACRR 402  
Db 371 QTPCTSGCYICRQSQOLEEQKQTVFGSEHRGQDDDDTTIKGRFNAINAVSMSCACPR 430  
Qy 403 SPRGLSPAHLGDCSSDLILIRKSRNFLRLHRTNQDQDFTFVYVRY 455  
Db 431 TPKGLSPAHLADGSAADILVRKCSRLDFLRHLHRTSNKQDFPFVYVRY 483  
RESULT 7  
Q9VNA6 PRELIMINARY; PRT; 687 AA.  
AC Q9VNA6  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE CG16708-PA (Cg16708-pb).  
GN ORFNames=CG16708;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
[1]  
SEQUENCE FROM N.A.  
RP MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfankoch C., Baldwin D.,  
RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Moberly C., Morris L., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nueskern D.R., Pacleeb J.M.,  
RA Palazolo M., Pittman K.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Massarman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster";  
RL Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426065; PubMed=12537568;  
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,  
RA Pacleeb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svirkas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila  
RL melanogaster euchromatic genome sequence";  
RN Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
RP [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426070; PubMed=12537573;  
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirkas R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celniker S.E.;  
RT "The transposable elements of the Drosophila melanogaster euchromatin:  
RT a genomics perspective";  
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Battencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
RT systematic review";  
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
RN [5]  
RP SEQUENCE FROM N.A.  
RG FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RG FlyBase;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AE003603; AAF52040.2; -;  
DR FlyBase; FBgn0037315; CG16708.  
DR GO: GO:0048102; P:autophagic cell death; IEP.  
DR GO: GO:0035071; P:salivary gland cell death; IEP.  
DR InterPro; IPR001206; DAGKC.  
DR Pfam; PF00781; DAGK cat; 1.  
DR SMART; SM00046; DAGKC; 1.  
SQ SEQUENCE 687 AA; 75690 MW; AAE081230A939412 CRC64;  
Query Match 27.8%; Score 801.5; DB 2; Length 687;  
Best Local Similarity 31.3%; Pred. No. 4.1e-58;  
Matches 197; Conservative 97; Mismatches 206; Indels 129; Gaps 16;  
Qy 11 QSVLWVQQRCAVSL--PARALLRWRSFGPCAGAPGACACVPVSEIITAVETDVHGK 68  
Db 59 QQLVWERLQIKOSPOGNEAKAPL-----PPDSPAPQGGICSYGPOSHVLHLD--DVVSI 111  
Qy 69 HQSGGKQWKMEK-----YAFTVHVCKRRAR-----HRWKWAQ 101  
Db 112 RSGDTKASSLKPPSPGSSRSCSGSDVAGCKPTSQYLITINYAMRLSKSDTCNRELRR 171  
Qy 102 VTFWCPPEQLCHLWLQTLREMLEKLT---SRPKHLLVFINPFGKGKGRIYKVPAPLF 158  
Db 172 LTFNSDPYIVRQWDQLRLHSSSPTRMVRRLVFINPYGGRKAGAQTYERHVRPIF 231  
Qy 159 TLASITTDIIIVTEHANOAKETLYEINIDKYDGVCGDGMFSEVLHGLIGRTORSAGVD 218  
Db 232 QLAGVADATCITTQRANQVLDLISHDLGVYDAVCCVGGDGTVAEVINGLIFROMRELGLD 291



```
Qy 271 VSVVH-----NSTLLRYSVSLGFGYGDIIKDSEKRWGLARY 311
Db 219 ISAVYSADAAKCEGASPGTGRPRQLLKFASALSYGYLGDIAVDSEKYRWGPKRY 278
Qy 312 DFSGLKTFSLHCHYEGTVSFLPAQHTVCSPRDRKPCRAGCFVCRQSKOOLEEEOKKALYG 371
Db 279 DYSGFKKFLANRGYNAEIVVHLDRRGKQDPNDGVRCLDKCARCKAKYGRDCGGERASY- 337
Qy 372 LEAAEDVEEMOVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDLILIRKCSRNFN 431
Db 338 ---EDDDTEPLVVRGKFLMVSGANISCSERSPOGSPSYCHLGDGLDLVLVRHTSMFN 394
Qy 432 LRFLIRHTNQ 441
Db 395 LRLLLTWSK 404

RESULT 10
Q6USK2 PRELIMINARY; PRT; 608 AA.
AC Q6USK2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Ceramide kinase.
GN Name=CERK;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22959668; PubMed=14563678; DOI=10.1101/gad.1140503;
RA Liang H., Yao N., Song J.T., Luo S., Lu H., Greenberg J.T.;
RT "Ceramides modulate programmed cell death in plants.";
RL Genes Dev. 17:2636-2641(2003).
DR EMBL; AY362552; AAQ62904.1; -.
DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
DR InterPro; IPR001206; DAGKC.
DR Pfam; PF00781; DAGK_cat; 1.
DR ProDom; PD005043; DAGKC; 1.
KW Kinase.
SQ SEQUENCE 608 AA; 68462 MW; 65EB9353692D9CA6 CRC64;

Query Match 22.4%; Score 648; DB 2; Length 608;
Best Local Similarity 31.6%; Pred. No. 2.5e-45;
Matches 182; Conservative 68; Mismatches 188; Indels 138; Gaps 19;

Qy 52 VPSVEIIAVEETDV-----HGKHGSGKQWKQKPYAFTVHCVKRARRH--RWK 98
Db 71 IKFSDIYAVEFVSGLVHSPKLGRLHAKCEFRLLNTQEMRYRTVHGFSQSPKEPCLWN 130
Qy 99 WAQVTFWCPBQLCHLWLTLEMLKLTSPKHLVFINPFGKGQKRIYKRVAPLF 158
Db 131 LAATFGHMDLQTCQWMDQLNYSLSIKEVERPRLNVFHPKSGKNGSKVWE-TVSKIF 189
Qy 159 TLAITTDIIIVTEHANQAKETLYEI---NIDKYDGIYCVCGDGMFSEVLHG-LIGRTQ-- 212
Db 190 IRAKYNKVIIVTERAGHAFDVMASIQNKELTYDGIITAVGGDGFNEITLGYLLSRLKVP 249
Qy 213 -----RSAG-----VDQ---NHPRAVLVPSL----- 231
Db 250 LPPSPDSFNSVQSGSSVPEPGDEVHTQKEHP---LLPSVQEVMMFRTVNGSCE 306
Qy 232 -----RIGIPAGSTDVCYSTGVGSDAFTSALHIVGDSLAMDVSSV--- 274
Db 307 GIEDPDHPFSERPRFGLIPAGSTDAIWMCTTGARDPVTSAHLIILGRKFLDAMQVVRW 366
Qy 275 --HNSNLT---LRYSVSLGFGYGDIIKDSEKRWGLARYDPSGLKTFSLHCHYEGTV 329
Db 367 KTASTSTIEPIRYAASFAGYGYGDVISESEKRYWMMGPKRYDYVGTGKIFLKRHSYEA 426
```

```
Qy 330 SFLPAQ-----HT-----VGSPRDRK-PCRAGCFVCRQSKOOLEEEOKKALYGL 373
Db 427 MFEAESENSKASLHTRSKTWPFRNTTTRSEKILCRANCKICNS-----KVGWNSASTTLN 481
Qy 374 AADVEQVQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDLILIRKCSRNFNPLR 433
Db 482 PCPEKTRCWTCTGRFLTSIGAAMVMSNRNERAPDGLVVD AHLSDGFLHLILIKDCSRPKYLW 541
Qy 434 FLIRHTNQ-ODQPDFTEVYVRVKFOFTSKHMEDESDLKEGKKRFGHICSSHPSCCC 492
Db 542 HLTELAKRGGEPLNFVEYHKTRAFITTS-----FG----- 573
Qy 493 TVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIE 528
Db 574 --EESVWNLDGEIFEAHQLSAQVLRLGLPLFASGPE 607
```

## RESULT 11

```
Q6H6H1 PRELIMINARY; PRT; 700 AA.
AC Q6H6H1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative ceramide kinase.
GN Name=P0519E06.23; Synonyms=QJ1003 B06.39;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, PAC
clone:P0519E06.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005006; BAD25678.1; -.
DR EMBL; AP004676; BAD25337.1; -.
DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
DR InterPro; IPR001206; DAGKC.
DR ProDom; PD005043; DAGKC; 1.
KW Kinase.
SQ SEQUENCE 700 AA; 77961 MW; FBB2C2A3DF31C75D CRC64;

Query Match 21.4%; Score 619; DB 2; Length 700;
Best Local Similarity 28.3%; Pred. No. 8.2e-43;
Matches 191; Conservative 86; Mismatches 192; Indels 206; Gaps 25;

Qy 1 MGATGAAPLQSVLWVQKQCAVSLEP-ARALLRWRSPPGAGAGADACSPVPSIIA 59
Db 70 LGQHGDIEEVSSSCW-----SSIIMQPKLESKLKF-----SDVAVVELLEVGP 112
Qy 60 VEE-----TDVHGKHGSGKQWKQKPYAFTVHCVKRARRHWK--AQVTFWCPBQLC 112
Db 113 VCEPWNARATVQGGKIN-----TEMNR---FVIHTVTRPRKRPSPWPVCEYIFGHKDOQTC 164
Qy 113 HLWLQTLREMLKLTSPKHLVFINPFGKGQKRIYKRVAPLFTLASITTDIIIVTEH 172
Db 165 KTWVEHKTCTNCKEQRKSLMVFVHPLCGKGRCKNWE-TVAPLFEAKVTKVIUTQR 223
Qy 173 ANQAKETLYEI---NIDKYDGIYCV----- 194
Db 224 AGHAYDTLASISDKDLKKFDGVIANTINACLSLFDIKHHYKMSARPENTLSYDPSAA 283
Qy 195 -----GGDGMFSEVL-----HGLIGRTORSA 215
```



```
Db 284 SGHKSMLFYCYFIINMKQEHNRNDLSNSELTDGDANAISGSSNTPPDDHEPLLSTTRST 343
Qy 216 GVD-----QHPRAVLV--PSS-LRIGIIPAGSTDC-----VCYSTVGTSDAE 255
Db 344 GLD1SSSDSDPCNQDQVPLVFPNNWFLGIIIPSGSDAIVLSPDVC-STTGERDPV 402
Qy 256 TSALHIVGDSLAMDVSV-----HNSLLAYSLSLLGFGYGDIIKDSKKRWLG 307
Db 403 TSALLIILGRRISLDIAQVVRKWSPSAEVSTVRYAASFAGYGFYGEVIRESEKRWNG 462
Qy 308 LARYDFSGLTKTFLSHHCYEGTVSFLPAQHT-----VGSPRDRK-CP 348
Db 463 PARYDFSGTWFLKHSYEAQVFLNGWTHSLTASAENNANGVQTLQVQNRHKTICR 522
Qy 349 ACCFVCR---QSKQLEBQKALYGLEAAEVEWQVVCVGFPLAINATNMSCACRRSPR 405
Db 523 TNLCLCKGTSTSEQSEDENPDS---SRTACETPKWVMSKGRFLSVGAIVSCRNERAPD 579
Qy 406 GLSPAHLGDSGLILIRKSRFNLRLIRHTNQ-----QDQDFTFVEYVRVKKPOF 460
Db 580 GLVADAHLSGDFLHLLLRDCP-----LPYLMHLTQFTYKKGSDPLSFVHHKTOAFTF 635
Qy 461 TSKHMEDESDLKEGKKRFGHCISHPSCCTVSNSSWNCDEGLVHSPAIEVRVHCOLV 520
Db 636 ISSHDE-----SYWNLDGELLQACEVSVQAFRGIV 665
Qy 521 RLFGARGIE--ENPKP 533
Db 666 NLFASGPEKMENTAP 680

RESULT 12
Q9LU45 PRELIMINARY; PRT; 533 AA.
AC Q9LU45;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similarity to unknown protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones."
RL DNA Res. 7:31-63(2000).
DR EMBL; AB023044; BAA97392.1; -.
DR GO; GO:0004143; P:diacylglycerol kinase activity; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
DR InterPro; IPR001206; DAGKC.
DR ProDom; PD005043; DAGKC; 1.
SQ SEQUENCE 533 AA; 60153 MW; 1B851C7606B03E0E CRC64;

Query Match 20.2%; Score 583.5; DB 2; Length 533;
Best Local Similarity 33.5%; Pred. No. 5.3e-40;
Matches 158; Conservative 57; Mismatches 157; Indels 99; Gaps 16;

Qy 52 VPVSEIIAETDV-----HGKHQSGKQKMEKPYAFTVCHVKARRH--RWK 98
Db 71 IKFSDIYAEVYSGLVHSPKLGRLHRAKECFRERLLNTQEMVRFVTHGFSQSPKPCLVN 130
Qy 99 MAQVTFWCPPEQLHLWLTQLEMLEKLTSRPKHLVFINPPGGKQGGKRIYERKVAFLF 158
Db 131 LAAFTFGHMDLQTCQSWMDQLNYSLSIKEVERPNLLVFFVHPKSGKNGSKYWE-TVSKIF 189
Qy 159 TLASITTDIIIVTEHANQAKETLVEI---NIDKDGIVCVGGDMFSEVLHG--LIGRTO-- 212
```

```
Db 190 IRAKNTKVIIVTERAGHAFDVMAIQNKELTYDGIITAVGGDFNEILNGYLLSRKVP 249
Qy 213 -----RSAG-----VDQ--NHPRAVLVPSSL----- 231
Db 250 LPPSPDSFNSVQSRGSSVPEPGDEVHETDQKEHP---LLPDSVQVWMNFRIEDPDHP 306
Qy 232 -----RIGIIPAGSTDCVYSTVGTSDAETSALHIVVGDSDLAMDVSV-----HNSL 280
Db 307 FSSERPRFGLIPAGSTDAIVMCTTGARDPVTSALHILGRKFLDAMDQVWRMKTASTSTI 366
Qy 281 ---LRYSVSLGFGYGDIIKDSKKRWLGARYDFSGLTKTFLSHHCYEGTVSFLPAQ-- 335
Db 367 EPIRYAASPAFYGYGDVISESEKYRMWGPRIYVGTGKIFLKHRSYEAEMFEEASE 426
Qy 336 -----HT-----VGSPRDRK-PCRAGCFVCRQSKQLEBQKALYGLEAAEVEW 381
Db 427 NSKASLHTRSTKTPFRNTTRSEKILCRANCKICNS-----KVGWNSATTTLNPCPEKTRW 481
Qy 382 QVCGCKFLAINATNMSCACRRSPRLGSPAHLGDSGLIILIRKSRFNFL 432
Db 482 CRTKGRFLSIGAAVMSNRNERAPDGLVVD AHLSDGFLHLILIKDCSRPKYL 532

RESULT 13
Q6UZF6 PRELIMINARY; PRT; 532 AA.
AC Q6UZF6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Ceramide kinase-like protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14681825;
RA Tuson M., Marfany G., Gonzalez-Duarte R.;
RT "Mutation of CERKL, a novel human ceramide kinase gene, causes
RT autosomal recessive retinitis pigmentosa (RP26).";
RL Am. J. Hum. Genet. 74:128-138(2004).
DR EMBL; AY357073; AARI3670.1; -.
DR GO; GO:0004143; P:diacylglycerol kinase activity; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
DR InterPro; IPR001206; DAGKC.
DR Pfam; PF00781; DAGK_cat; 1.
DR ProDom; PD005043; DAGKC; 1.
KW Kinase.
SQ SEQUENCE 532 AA; 59602 MW; C73B590F7C25EED1 CRC64;

Query Match 20.0%; Score 576.5; DB 2; Length 532;
Best Local Similarity 27.6%; Pred. No. 2.1e-39;
Matches 149; Conservative 102; Mismatches 195; Indels 93; Gaps 12;

Qy 29 RALLRWRSPPGAPGACADACSVPSSE-----IIAVEE----- 62
Db 45 RILLR-----GIFEIGRSDCVLSEALRWPIQPERPAGDSKYDLLCKEFLDKD 97
Qy 63 -TDVGHQSGKQKMEKPYAFTVCHVKARRHWRKVAQVTFWCPEQLCHLWLQTLURE 121
Db 98 IFSVKLRCSRCSVKQQRSGTLGITLFLICLKEQKLNKSTLDLNLSDHDCDIWFRQPK 157
Qy 122 MLEKLTSRPKHLVFINPPGGKQGGKRIYERKVAFLFTLASITTDIIIVTEHANQAKETLY 181
Db 158 ILAGFNRPRKSLTLLNPQSHKKEATQVYEKVEPLKLAKIGIKTDVTIMBEGHALSLK 217
Qy 182 BINIDKDGIVCVGGDMFSEVLHGLIGRTQSRAGVQDNHPRAVLVP--SSLRIGIIPAG 239
Db 218 ECELOGDFGVVCGDGSASEVAHALLLRQAKNAGMETDR---ILTVPRAQLPLGLIPAG 274
```



Qy	240	STDCVCYSTGCTSDAETSAIHITVCDLSAMDVSSVHNSHTLLRYSSVSLGGYGPYGDIIKD	299
Db	275	STNVLAHS LHGVPVHITATLHIMGHVQLVDCTFTAGKLLRFGFSAM- FPGSGRTAL	333
Qy	300	SEKRWLG- LARYDFSGLKTLFSLHCHYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCRQSK	358
Db	334	AEKYRWMSPNQRDFAVVKALAKAEDCEISFLFPNSS	372
Qy	359	QOLEBEOKALVGLEAADVEWQVCKFLAIATNMWSCARRSPGSLSPAHLGDGSS	418
Db	373	---DVOERRAAGSPKDCNDQWQMIQQGLNVSIMAIPLCSVAPRGLAPNRLNNGSM	429
Qy	419	DLILRKCSRNFRLFRHRTNODQDQFTFVVVVRVKKFTQTSKH-----MEDESD	471
Db	430	ALLIAENTSRPFIKHLRYASVKQNFPFVEITYTVEEVKVPNNVTGCGYNEFEDEET	489
Qy	472	LKEGGKKRFGHTCSSHPSCCCTTVSNSSWNCDEVLH- SPAIEVRVHCOLVRLFARGIEE	529
Db	490	ASE-----NC-----FPWNVDGLMEVASEVHRLHLPRLTISLGVGSMEE	528

RESULT 14	
Q9TZ11	
ID	Q9TZ11
AC	PRELIMINARY;
DT	PRT; 549 AA.
DT	Q9TZ11;
DT	01-MAY-2000 (TrEMBLrel. 13, Created)
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE	Hypothetical protein T10B11.2.
GN	Names=T10B11.2; ORFNames=T10B11.2;
OS	Caenorhabditis elegans.
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC	Rhabditidae; Feloderinae; Caenorhabditis.
OX	NCBI TaxID=6239;

RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RT WormBase Consortium;  
RG "Genome sequence of the nematode *C. elegans*: a platform for  
RT investigating biology. The *C. elegans* Sequencing Consortium.";  
RL Science 282:2013-2018(1998).

RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Minx P., Kemp K.;  
RT "The sequence of *C. elegans* cosmid T10B11.";  
RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.

KN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.

RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Wilson R.;  
 RL Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.

RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RG WormBase Consortium;  
RL Submitted (SEP-2004) to the ENBL/GenBank/DDJF databases.

DR *EMBL*, *AF089353*; *AF089400.1*,  
DR *PIR*, *T33517*, *T33517*.  
DR *WormBase*; *WBGene0020398*; *T10B11.2*.  
DR *WormPep*; *T10B11.2*; *CE18241*.  
DR *InterPro*; *IPR001206*; *DAGKC*.  
DR

Query Match 19.7%; Score 567.5; DB 2; Length

Query Match

[illegible]

## RESULT 15

Q949C3	PRELIMINARY;	PRT;	586 AA.
ID	Q949C3		
AC	Q949C3;		
DT	01-DEC-2001 (TReMBLrel. 19, Created)		
DT	01-DEC-2001 (TReMBLrel. 19, Last sequence update)		
DT	01-MAR-2004 (TReMBLrel. 26, Last annotation update)		
DE	Hypochemical protein C875ERIPDS.		
GN	Name=C875ERIPDS;		
OS	Oryza sativa (Rice).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	EC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
OC	Ehrhartioideae; Oryzaceae; Oryza.		
OX	NCBI_TaxID=4530;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21329048; PubMed=11435398; DOI=10.1101/gr.GR-1617R;		
RA	Mayer K., Murphy G., Tarchini R., Wambutt R., Volckaert G., Pohl T.		
RA	Dueterhoeft A., Stiekema W., Entian K.D., Terry N., Lemcke K.,		
RA	Haese D., Hall C.R., van Dodeeweerd A.M., Tinge S.V., Mewes H.W.,		
RA	Bavan M., Bancroft I.;		
RT	"Conservation of microstructure between a sequenced region of the		
RT	genome of rice and multiple segments of the genome of Arabidopsis		
RL	thaliana.";		
RL	Genome Res. 11:1167-1174 (2001).		
DR	EMBL; AJ307662; CAC39069.1; -.		
DR	Gramene; Q949C3; -.		
DR	GO; GO:0004143; F-diacylglycerol kinase activity; IEA.		
DR	GO; GO:0007205; P-protein kinase C activation; IEA.		
DR	InterPro; IPR001206; DAGKC.		
DR	ProDom; PD005043; DAGKC; 1.		
DR	Hypochemical protein.		
KW	SEQUENCE 586 AA; 64892 MW; 1C45DFBB670E1E22 CRC64;		
SO	SEQUENCE		

Query Match 16.6%; Score 480; DB 2; Length 586;  
Best Local Similarity 26.8%; Pred. No. 2.8e-31;  
Matches 152; Conservative 81; Mismatches 173; Indels 162; Gaps 21;



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 6, 2005, 07:01:07 ; Search time 5163.87 Seconds  
(without alignments)  
3958.373 Million cell updates/sec

Title: US-10-631-958-10

Perfect score: 2888

Sequence: 1 MGATGAAPLQSLVWVKQOR.....QLVRLFARGIENPKPDHS 537

Scoring table:

```
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
```

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US10631958/runat.02092005.165816.4562/app_query.fasta_1.1941
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODES=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10631958 @CGN 1.1 9235 @runat.02092005.165816.4562 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hsc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2433.5	84.3	1824	3 AK042077	Mus muscu
2	2424.5	84.0	4248	3 AK052269	Mus muscu
3	1518	52.6	1063	4 BM479389	AGENCOURT
4	1367	47.3	1059	5 BQ054406	AGENCOURT
5	1321	45.7	1078	5 BQ063738	AGENCOURT
6	1299.5	45.0	797	7 CK000755	AGENCOURT
7	1284	44.5	1047	5 BQ057191	AGENCOURT
8	1275	44.1	713	7 CK000525	AGENCOURT
9	1260	43.6	820	6 CD855311	AGENCOURT

10	1244	43.1	732	7 CF135528	UI-HF-BNO
11	1205	41.7	758	6 CB246749	UI-M-FIO-
12	1181	40.9	653	5 BX952302	DKF2p781L
13	1170	40.5	661	7 CN296312	170005321
14	1053.5	36.5	812	7 CK603033	AGENCOURT
15	1012	35.0	581	5 BP224560	BP224560
16	1003	34.7	1017	5 BQ879739	AGENCOURT
17	997	34.5	584	5 BP310011	BP310011
18	986	34.1	779	5 BU265377	603374309
19	982	34.0	584	5 BP309990	BP309990
20	971	33.6	763	7 CK364327	AGENCOURT
21	968	33.5	573	7 CF138275	UI-HF-BNO
22	964	33.4	819	7 CN235868	WLB086C10
23	959	33.2	550	7 CF138634	UI-HF-BNO
24	943.5	32.7	759	5 BU264162	603814452
25	942	32.6	541	7 CN296311	170005321
26	925	32.0	725	5 BU214294	603756937
27	924	32.0	607	6 CB581157	AGNNUC:N
28	916	31.7	576	6 CB608292	AMGNNUC:N
29	916	31.7	582	5 BP274786	BP274786
30	914	31.6	653	6 BY754042	BY754042
31	909	31.5	608	6 CB580936	AMGNNUC:N
32	876	30.3	553	6 CA578972	K0731A08-
33	873	30.2	674	1 AJ739034	AJ739034
34	871	30.2	550	2 AW499858	UI-HF-BNO
35	849	29.4	902	5 BU73554	603811294
36	838	29.0	545	6 CB611947	AMGNNUC:N
37	829	28.7	662	2 BB638867	BB638867
38	828	28.7	488	5 BU430459	UI-HF-BNO
39	813	28.2	540	6 CB612897	AMGNNUC:N
40	811	28.1	474	2 AW503999	UI-HF-BNO
41	806	27.9	638	1 AJ739022	AJ739022
42	805	27.9	449	2 AW503483	UI-HF-BNO
43	797.5	27.6	759	5 BU290807	604162346
44	796	27.6	656	6 BY739924	BY739924
45	786.5	27.2	680	2 BB235873	BB235873

#### ALIGNMENTS

RESULT 1	AK042077	1824 bp	linear	HTC 03-APR-2004
LOCUS	AK042077	1824 bp	musculus 3 days neonate thymus cDNA, RIKEN full-length enriched library, clone:A630056D11 product:DA59H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS) (FRAGMENT) homolog [Homo sapiens], full insert sequence.	
DEFINITION	AK042077	GI:26334912	HTC; CAP trapper.	
ACCESSION	AK042077	GI:26334912	Mus musculus	
VERSION	AK042077	GI:26334912	Mus musculus (house mouse)	
KEYWORDS	HTC; CAP trapper.		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
SOURCE	Mus musculus			
ORGANISM	Mus musculus			
REFERENCE	1			
AUTHORS	Carninci, P. and Hayashizaki, Y.			
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)			
MEDLINE	99279253			
PUBMED	10349636			
REFERENCE	2			
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)			
MEDLINE	20499374			
PUBMED	11042159			
REFERENCE	3			
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,			

Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multipillar sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)  
 20530913  
 11076861

#### REFERENCE

#### AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the  
 FANTOM Consortium.

#### TITLE

#### JOURNAL

#### REFERENCE

#### AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research  
 Group Phase I & II Team.

#### TITLE

#### JOURNAL

#### REFERENCE

#### AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
 Fukuda, S., Furuno, M., Hangaki, T., Hara, A., Hashizume, W.,  
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
 Okazaki, Y., Saito, R., Saichoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,  
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
 Muramatsu, M. and Hayashizaki, Y.

#### TITLE

#### JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
 Physical and Chemical Research (RIKEN), Laboratory for Genome  
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,  
 URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,  
 Fax: 81-45-503-9216)

#### COMMENT

CDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics  
 Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome  
 Trust/MRC building Addenbrookes Hospital Cambridge) whose  
 assistance we gratefully acknowledge.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/.

#### FEATURES

#### source

1. 1824  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM,DB:A630056D11"  
 /db\_xref="taxon:10090"  
 /clone="A630056D11"  
 /tissue\_type="thymus"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="3 days neonate"  
 85. 1680

#### CDS

note="unnamed protein product; DA59H18.2 (NOVEL PROTEIN  
 SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED)  
 PROTEINS) (FRAGMENT) homolog [Homo sapiens] (SPTFQ9UGES,  
 evidence: FASTV, 86.5%ID, 100%length, match=945)  
 putative"

/codon\_start=1

/protein\_id="BAC31157.1"

/db\_xref="GI:26334913"

/translation="MGAMGAAEPLHSLVWVKRRCAVSLSEPARALLRWRSPEPGPSA

#### ORIGIN

Alignment Scores: Length: 1824  
 Pred. No.: 8,32e-231 Matches: 453  
 Score: 2433.50  
 Percent Similarity: 90.19% Conservative: 34  
 Best Local Similarity: 83.89% Mismatches: 48  
 Query Match: 84.26% Indels: 5  
 DB: 3 Gaps: 1  
 US-10-631-958-10 (1-537) x AK042077 (1-1824)  
 Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrrpVallysGlnGlnArg 20  
 Db 85 ATGGGGCAATGGGGCGCGGAGCCGCTGCCTGCTGTGGTGAACACGGCAGCGC 144  
 Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrrpArgSerProGlyPro 40  
 Db 145 TGTGCGCTCAGCTGGAGCCGCGGGGCTCTGCTACGCTGGTGGCGGAGCCCGGAGCC 204  
 Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal 60  
 Db 205 GGGCCCTCCGCGCGGCTGCATGCCGCTCGGTACTGCTGTCGAGATCATCGCTGTT 264  
 Qy 61 GluLthrAspValHisGlyLysGlnGlySerGlyLysTrrpGlnLysMetGluLys 80  
 Db 265 GAGGAAAACAGCAGCTCGGAGAACACGCTCCAGTGGCGATGGCTAAATGGAAT 324  
 Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrrpLysTrrpAla 100  
 Db 325 CCGTTCGATTCAGTCCACCGTGTGAGGAGTACGACATCACCGCTGGAAGTGGCA 384  
 Qy 101 GlnValThrPheTrrpCysProGluGlnGlnCysHisLysLeuTrrpLeuLysArg 120  
 Db 385 CGGTGACCTTCTGGAGCGCGCAGCAGCTGTGTCCCTGTGCTGCAGACCTCCGT 444  
 Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheLeuAspPhe 140  
 Db 445 GGGCTGTGGAGACCTTGACTTCAAGACCGAGCATTGCTGGTATTTCATCAACCTTC 504  
 Qy 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160  
 Db 505 GGAGGAAAGGTGAGGCAAGCGCATCTATGAAAACACAGTGGCGCTCTGTTTACCTTG 564  
 Qy 161 AlaSerIleThrAspIleIleValThrGluHisAlaAlaGlnAlaLysGluThrLeu 180  
 Db 565 GCTTCCATCACTACGAGATCATATTACAGAGCATGCCAACCAAGCAAGCAGACATTTA 624  
 Qy 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200  
 Db 625 TACGAGATCAACACAGACAGCATATGATGGCATCGTGTGCTGAGTGGGCGGATGTTT 684  
 Qy 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220  
 Db 685 AGCGAGGTGTCATGGGTGATTGGGAGGAGCGCAGCAGAGCGCTGGTATCGACCCCAAT 744  
 Qy 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240  
 Db 745 CACCCCGAGCGGTGCTGGTCCAGTACCCCTAGGATCGGCATCATACCCGAGGGTCC 804  
 Qy 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260  
 Db 805 ACAGATTGTGTGTTTACTCAACAGTGGGCACAAACGACCGCAGAGACATCGGCTTTGCAC 864



```
/clone_lib="RIKEN full-length enriched mouse cDNA library"  
/dev_stage="13 days embryo"  
260_-1855  
/note="unnamed protein product; DA59H18.2 (NOVEL PROTEIN  
SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED)  
PROTEINS) (FRAGMENT) homolog [Homo sapiens] (SPTP[Q9UGES,  
evidence: FASTY, 86.5%ID, 100%length, match=945)  
putative"  
/codon_start=1  
/protein_id="BAC34908.1"  
/db_xref="GI:26342492"  
/translation="MGAMGAEEPLHSLVWVRRRCVSLSEPARALLRWRSPBGPSPA  
PGADARSLVSEIIAVEKDDCEKHAAGSRWKMENPFAFVHRVKRVRHRRWKWARV  
TFWADQLCHLWLOTURGLLELVSFSPKLLVLPNFGKGGQGRKYETKTVAPLPTL  
ASITTEIIITHEANAQKETLYEINTDSYDGI VCVGGDGMFSEVLHGVIGRTQOSAGID  
PNHPRVLVSTLRIGLIPAGSDICVSVTGNDAETSAHIIIGDSLDAIDVSSVHY  
HTTLIRYSVLLGFGVDLIIKOSKKRWMLGVDFSGUKTFLSHQYISGTLISFLPA  
QHTVGSPRDKRCRAGFVCRSQKQLEEEKKALYGLSENAEEMEQVTCGKFLAIN  
ATNNSCACPRS PGLSLFPALHGDGSDLIILIRKCSRFNFLRIHTNQDQDFTFV  
EVRVKCFHTSKHVEVDNDSKEQEKQKFKICKDRPSC TCSASRSSWNCDEGVHHS  
PAIEVRVHCQLVRLFARGIEEBS"
```

CDS

## ORIGIN

## Alignment Scores:

Pred. No.:	2,42e-229	Length:	4248
Score:	2424.50	Matches:	452
Percent Similarity:	90.00%	Conservative:	34
Best Local Similarity:	83.70%	Mismatches:	49
Query Match:	83.95%	Indels:	5
DB:	3	Gaps:	1

US-10-631-958-10 (1-537) x AK052269 (1-4248)

Qy	1	MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnArg	20
Db	260	ATGGGGGCNAATGGGGGGCGGGAGCCGCTGCACCTCCGTGTGGGTGAACCGCGCAGC	319
Qy	21	CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro	40
Db	320	TGTGCGGTACAGCTGGAGCCGCGGGGCTCTGTACGCTGGTGGCGGAGCCCGGAGCC	379
Qy	41	GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal	60
Db	380	GGGCCCCCTCGCGGGTGGCGATGCCCGCTCGGTCTGGTGTGGAGATCATCGCTGT	439
Qy	61	GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys	80
Db	440	GAGGAAAAAGACGATCGAGAAACACGCTCCAGTGGCGCGATGGCATAAATGGAAAT	499
Qy	81	ProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrpLysTrpAla	100
Db	500	CCGTTCGATTCACAGTCCACCGTGTGAAGCGAGTACGACATCACCGCTGGAAGTGGCA	559
Qy	101	GlnValThrPheTrpCysProGluGlnLeuGlnCysHisLeuTrpLeuGlnThrLeuArg	120
Db	560	CGGCTGACCTTCTGGAGCCCGGAGCAGCAGTGTGTCACTGTGGCTGCAGACCTCCGT	619
Qy	121	GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe	140
Db	620	GGGTCTGTGGAGAGCCCTGACTTCAAGACCGACCTGTGCTGGTATTATCAACCCCTTC	679
Qy	141	GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu	160
Db	680	GGAGGAAAGGTTCAGGCAACGGCATCTATGAAAAACAGTGGCGCTCTGTTTACCTTG	739
Qy	161	AlaSerIleThrAspIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu	180
Db	740	GCTTCCATCACTACGGAGATCATATTACAGAGCATGCCAACCAAGCAAGGAGACTTA	799
Qy	181	TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe	200
Db	800	TACGAGATCAACACAGACAGCATATGATGGCATGTGTGCTGCTAGTGGGAGCGGCATGTC	859

Qy	201	SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn	220
Db	860	AGCGAGGTGCTGATGGGGTGATTGGAGGACGACGAGAGCGCTGTTATCGACCCCAAT	919
Qy	221	HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer	240
Db	920	CACCCCGGACCGTGTGGTGGCCAGTACCCTCAGGATCGCATATACCCGCGAGGCTCC	979
Qy	241	ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis	260
Db	980	ACAGATTGTGTGTACTCAACAGTGGGCACAAACGACGACGACGATCGCTTTGCAC	1039
Qy	261	IleValValGlyAspSerLeuAlaMetAspValSerSerValHisAsnSerThrLeu	280
Db	1040	ATCATATTGGGGACTCACCTGGCAATAGACGTGCTCTCTGTGCATACCATATAACGCTG	1099
Qy	281	LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer	300
Db	1100	CTGCGGTACTCGGTTTCTCTGTGGGTACGGTTTCTACGGGACTTAATCAAGGACAGT	1159
Qy	301	GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu	320
Db	1160	GAAGAAGAACGGTGGATGGGCCCTCGTCCGATATGATTTCTCAGGTTGAAGACCTTCTC	1219
Qy	321	SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer	340
Db	1220	TCTCATCAGTACTATGAAGGACACACTGTCTCTCTCCACGACAGACACGGTGGATCT	1279
Qy	341	ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln	360
Db	1280	CCACGGGACAATAAACCCCTCCCGGCTGGGTGTCTGTGTGCAGGCGACAGCAACAG	1339
Qy	361	LeuGluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu	380
Db	1340	CTGAAGAAGAAAGAAAGCCCTGTATGGCTGGAGAACGCCGAGGAAATGAAGAGAG	1399
Qy	381	TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys	400
Db	1400	TGCAGTGCATGTGGGAAGTTCTTGGCCATCAATGCCACCAACATGTCTGTGCTGT	1459
Qy	401	ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu	420
Db	1460	CCTCGGAGCCCTGGGGCCCTGCTCCCATTTGGCCATCTGGGAGATGGGTCTTCTGACCTC	1519
Qy	421	IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn	440
Db	1520	ATCTTATCCGGAAGTCTCCAGTTCACTTCTGAGATTCCTCATCCGCGCACACCAAC	1579
Qy	441	GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe	460
Db	1580	CAGGAGGACCAGTTCGACTTCACTTCTTGAAGTTTATCGAGTCAAGAAATTCACCTTC	1639
Qy	461	ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe	480
Db	1640	ACGTGGAAGCACCTGGAAATACGAGGACAATGATCTGAAGGAACAAGAGAGCAAGTTT	1699
Qy	481	GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTyrAsn	500
Db	1700	GGGAAGATCTGCAAGGACACACCTCTTGTCACTTGTCTAGCTCCAGCCTCCGGAAC	1759
Qy	501	CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal	520
Db	1760	TGCGATGGCAAGTCAATGACACAGCCCGGCAATGAGTCAAGGTCCCATCTGCAGCTGGT	1819
Qy	521	ArgLeuPheAlaArgGlyIleGluGlu-----AsnProLysProAspSer	535
Db	1820	CGCTCTTCTCGGGGAATCGAGGAAGAGTCAATAGCAAGAAACCCCAAGCCCGAGGC	1879

RESULT 3

BM479389

LOCUS

DEFINITION

BM479389 1063 bp mRNA linear EST 05-PEB-2002  
AGENCOURT\_6418742 NIH\_MGC\_67 Homo sapiens cDNA clone IMAGE:550281  
5', mRNA sequence.



[illegible]

```

/db_xref="taxon:9606"
/clone="IMAGE:5803668"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/notes="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

```

## ORIGIN

```

Alignment Scores:
Pred. No.:      5.63e-125      Length:      1059
Score:          1367.00      Matches:      276
Percent Similarity: 92.74%      Conservative: 5
Best Local Similarity: 91.09%      Mismatches: 14
Query Match:     47.33%      Indels:      8
DB:              5          Gaps:      4

```

US-10-631-958-10 (1-537) x BQ054406 (1-1059)

```

Qy 133 LeuLeuValPheIleAsnProPheGlyGlyGlyGlnGlyLysArgIleTyrGluArg 152
Db 15 TTACTGTATTATCAACCCGTTTGGAGAAAGACAAAGCAGCGGATATATGAAGA 74
Qy 153 LysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleValThrGluHis 172
Db 75 AAAGTGGCACCACCTGTTCCACCTTAGCTCCCATCCACTGACATCATCGTTACTGAACAT 134
Qy 173 AlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleVal 192
Db 135 GCTAATACGGCCAAAGGAGACTCTGTATGAGATTAACTAGACAAATACGACGGCATCGTC 194
Qy 193 CysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGln 212
Db 195 TGTGTGGCGGAGATGATGATGTTTCAAGCAGGTGTCGACGGTCTGATTTGGAGACCGAG 254
Qy 213 ArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArg 232
Db 255 AGGAGCGCGGGTCCGACAGAACACACCCCGGGCTGTGCTGGTCCCGTAGCTCCGG 314
Qy 233 IleGlyIleIleProIleAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSer 252
Db 315 ATTGGAAATCATTCGCGAGGGTCAACGGACTGCGGTGTGTTACTCCACCGTGGGCACGAGC 374
Qy 253 AspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSer 272
Db 375 GACGAGNAACCTCGGCGCTGCATATCGTTGTTGGGACCTCGCTGGCCATGGATGTCTCC 434
Qy 273 SerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPhe 292
Db 435 TCAGTCCACCAACAACAGCACACTCTTCGCTACTCCGGTCCCTGCTGGGCTACGGCTTC 494
Qy 293 TyrGlyAspIleIleLysAspSerGluLysIysArgTrpLeuGlyLeuAlaArgTyrAsp 312
Db 495 TACGGGGACATCATCAAGGACAGTGAGAAAGAAACGGTGGTGGTCTTGGCAGATACGAC 554
Qy 313 PheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeu 332
Db 555 TTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATATGACCGACAGTGTCTTCTCTC 614
Qy 333 ProIleGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPhe 352
Db 615 CCTGCACAACAACAACGGTGGGATCTCCAAGGATAGGAAGCCCTCCCGGGCAGGATGCTTT 674
Qy 353 ValCysArgGlnSerLysGlnGlnLeuGluGlnLysAlaLeuTyrGlyLeu 372
Db 675 GTTTGCAGGCAAGCAAGCAGCAGCTGTGAGGAGGACAGAAAGCAAGCACTGTATGGGTG 734

```

```

Qy 373 GluAlaAlaGluAspValGluGluTrpGlnValValCys-GlyLysPheLeuAlaIleAs 392
Db 735 GAAACTCGCGAGGACGCTGGAGGAATGCAAGTCGCTGTGGAAAGTTCCTGGCCATTAA 794
Qy 392 nAlaThr-AsnMetSerCysAla---CysArgArgSerProArgGlyLeuSerProAlaA 411
Db 795 TGCCACAAAACATGCCCTGTGCTTGTCCGCCCGGACCTCCAGGGCCCTCTCCCCCGG 854
Qy 411 la-----HisLeuGly---AspGlySerSerAspLeuLeuLeuLeuLeuLeuLeu 426
Db 855 GCTGGCCCAACTTGGGAAGACGGGGTCTTCCGACCTCTCTTCTTCTTCTTCTTCTT 914
Qy 426 YsSer 427
Db 915 GCTCC 919

```

```

RESULT 5
BQ063738 1078 bp mRNA linear EST 02-APR-2002
LOCUS
DEFINITION AGENCOURT 6873251 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5925382
5', mRNA sequence.
ACCESSION BQ063738
VERSION BQ063738.1 GI:19891754
KEYWORDS EST.
SOURCE Homo sapiens (human)

```

```

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1078)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-foemail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

```

```

Plate: LICM2096 row: p column: 23
High quality sequence stop: 640.

```

## FEATURES

```

source
1..1078
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5925382"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/notes="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

```

## ORIGIN

```

Alignment Scores:
Pred. No.:      2.2e-120      Length:      1078
Score:          1321.00      Matches:      260
Percent Similarity: 92.31%      Conservative: 4
Best Local Similarity: 90.91%      Mismatches: 15
Query Match:     45.74%      Indels:      7
DB:              5          Gaps:      4

```

US-10-631-958-10 (1-537) x BQ063738 (1-1078)

```
Qy 90 LysArgAlaArgHisArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGlu 109
Db 1 AAGAGACGACGACGCGCGTGGAGTGGCGCGAGGTGACTTTCTGGTGTCCAGAGGAG 60

Qy 110 GlnLeuCysHisLeuTrpLeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArg 129
Db 61 CAGCTGTGTCACTTGTGGCTCGAGACCCCTCGGGAGATGCTGGAGAGCTGACGTCCAGA 120

Qy 130 ProLysHisLeuValPheIleAsnProPheGlyGlyGlyGlnGlyLysArgIle 149
Db 121 CCAAGACATTTACTGGTATTTATCAACCGTTTGGAGGAAAGGCAAGGCAAGCGGATA 180

Qy 150 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleVal 169
Db 181 TATGAAGAGAAAGTGGCCACTGTTCACCTTAGCCTCCATCACCACCTGACATCATCGTT 240

Qy 170 ThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAsp 189
Db 241 ACTGAACATGCTTAATCAGGCCAAGAGACTCTGTATGAGATTAAATAGACAAATACGAC 300

Qy 190 GlyIleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGly 209
Db 301 GGCATCGTCTGTGTGGCGGAGATGGTATGTTTCAGCGAGGTGCTGCACGGTCTGATTGG 360

Qy 210 ArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSer 229
Db 361 AGGACGCGAGAGAGCGCGGGTTCGACAGAACCCACCCCGCGGTGTGCTGCTCCCGAGT 420

Qy 230 SerLeuArgIleGlyIleProAlaGlySerThrAspCysValCysTyrSerThrVal 249
Db 421 AGCTCCGGATTGGAAATCATTCGCCGAGGTACAGGACTGGGTGTGTTACTTCCACCGTG 480

Qy 250 GlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMet 269
Db 481 GGCACCGAGCGAGAAACCTCGCGCTGCATATCGTTGTTGGGAGCTCGTGGCCATG 540

Qy 270 AspValSerSerValHisAsnSerThrLeuLeuArgTyrTyrSerValSerLeuLeuGly 289
Db 541 GATGTGTCTCAGTCACACACAGCACACTCTCGCTACTCTGCTGCTGCTGCTGGGC 600

Qy 290 TyrGlyPheTyrGlyAspIleIleLysAspSerGluLysLysAspTrpLeuGlyLeuAla 309
Db 601 TACGGCTCTTACGGGAGCATCATCAAGAGAGTGAAGAAACGGTGTGGGTCTTGCC 660

Qy 310 ArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrVal 329
Db 661 AGATAGGACTTTTTCAGGTTTAAAGACTTCTCTCCACCACTGCTATGAAGAGACAGG 720

Qy 330 SerPheLeuProAlaGlnHisThrValGlySer---ProArgAsp-ArgLysProCysAr 348
Db 721 TCCTTCTCTCCCTGCACACACACGCGGTGGGATCTCCAGGGAATAAGGAAGCCCTGCC 780

Qy 348 gAla---GlyCys-----PheValCysArgGlnSerLysGlnGlnLeuGluGlu--- 364
Db 781 GGCAGCAGGATGCCTTTGTTGGCCAGCCCAAGCAAGCCAGCCAGCTTGGAAAGCGAGG 840

Qy 365 ----GlnLysLysAla 368
Db 841 AGCCCCAAGAAAGCC 856

RESULT 6
CK000755
LOCUS
DEFINITION AGNCOURT 16363467 NIH_MGC_221 Homo sapiens cDNA clone
IMAGE:30707875 5', mRNA sequence.
ACCESSION CK000755
VERSION CK000755.1 GI:38526789
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 797)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: James Martin, University of Iowa  
cDNA Library Preparation: M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
plate: NDAM1073 row: c column: 20  
High quality sequence stop: 656.

FEATURES  
source

Location/Qualifiers  
1..797

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30707875"  
/lab\_host="DH10B Tona"  
/clone\_lib="NIH\_MGC\_221"  
/note="Organ: mixed; Vector: pYX-Asc; Site 1: EcoRI;  
Site 2: NotI; Library is oligo-dT primed and directionally  
cloned. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT  
primer containing a Not I site. Double strand cDNA was  
size selected according to mRNa size fraction, ligated with  
EcoR I adaptor, digested with Not I and then cloned  
directionally into pYX-Asc vector. Average insert size  
4-5kb. Adaptors 5' (AATTCGGCAGCAGG)3' and 5'd  
(CCTCGCGCG)3'. 3' linker sequence - GCGGCGCTGAGGCC T18.  
Sequencing primers 3' end: T3 promoter primer 5'd  
(ATTACCCCTCCTAAAGGGA)3'. 5' End: T7 promoter primer 5'd  
(TAATACGACTCACTATAGG)3'. Library was constructed in the  
laboratory of M. Bento Soares. Note: this is a NIH\_MGC  
Library"

ORIGIN

Alignment Scores:  
Pred. No.: 1,91e-118 Length: 797  
Score: 1299.50 Matches: 250  
Percent Similarity: 96.54% Conservative: 1  
Best Local Similarity: 96.15% Mismatches: 8  
Query Match: 45.00% Indels: 1  
DB: 7 Gaps: 1

US-10-631-958-10 (1-537) x CK000755 (1-797)

Qy 152 ArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGlu 171  
Db 9 CGGCAGCAGGACACACCTGTTTCCCTTAGCTCCATCACCCTGACATCATCGTTACTGAA 68  
Qy 172 HisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIle 191  
Db 69 CATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGACGCGATC 128  
Qy 192 ValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThr 211  
Db 129 GTCTGTGTGCGCGGAGATGGTATGTTTACGCGAGGTGCTGCACGGTCTGATTGGGAGGAC 188  
Qy 212 GlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeu 231  
Db 189 CAGAGGAGCGCGGGGTGCGACCAACACCCCGGGCTGTGCTGGTCCCCAGTAGCCTC 248  
Qy 232 ArgIleGlyIleIleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThr 251  
Db 249 CGGATTGGATCATTTCCCGCAGGGTCAACGAGACTGCGGTGTGTACTTCCACCGTGGGCACC 308

Qy 252 SerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspVal 271  
 Db 309 AGCAGCAGAAACCTCGGCGCTCATATGCTTGTGGGACTCGCTGGCATGGATGTG 368  
 Qy 272 SerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGly 291  
 Db 369 TCCTCAGTCCACCAACACAGCACACTCTCTCGCTACTCCGCTGCTCGCTGGGTACGGC 428  
 Qy 292 PheTyrGlyAspIleLeuLysAspSerGluLysValArgTrpLeuGlyLeuAlaArgTyr 311  
 Db 429 TTCTACGGGACATCATCAGGACAGTGAAGAAGCGTGTGGGTCTTGCCAGATAC 488  
 Qy 312 AspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPhe 331  
 Db 489 GACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTC 548  
 Qy 332 LeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCys 351  
 Db 549 CTCCTTCGACAAACACCGTGGGATCTCCAAGGATAGGAAGCCCTCCCGGCGAGGATGC 608  
 Qy 352 PheValCysArgGlnSerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGly 371  
 Db 609 TTTGTTTTCAGGCAAGCAAGCACAGCTGGAGGAGGACGAAGAAGCACTGTATGGT 668  
 Qy 372 LeuGluAlaAlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIle 391  
 Db 669 TTGGAAGCTCGGAGGACGTGGAGAGTGGNCAGTCTGCTGGGAAAGTTTCTGGGCCNA 728  
 Qy 392 AsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAla 411  
 Db 729 TCATGCAAAACATGTCTGTGCTTGTGCGCGAGGCCA---GGGCTCTCCCGGCTGCC 785

RESULT 7  
 BQ057191  
 LOCUS BQ057191 1047 bp mRNA linear EST 29-MAR-2002  
 DEFINITION AGENCOURT\_6769595 NIH\_MGC\_99 Homo sapiens cDNA clone IMAGE:5812382  
 5', mRNA sequence.  
 ACCESSION BQ057191  
 VERSION BQ057191.1 GI:19816531  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1047)  
 NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Lou Staudt  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LICM2062 row: 1 column: 15  
 High quality sequence stop: 535.  
 Location/Qualifiers  
 1. 1047  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5812382"  
 /tissue\_type="lymphoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_99"  
 /note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGACGAG(G). Size-selected >500bp for average insert size  
 1.8kb. Library constructed by Ling Hong in the laboratory

of Gerald M. Rubin (University of California, Berkeley)  
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
 II RT (Life Technologies). Note: this is a NIH\_MGC  
 Library."

ORIGIN  
 Alignment Scores: Pred. No.: 1,02e-116 Length: 1047  
 Score: 1284.00 Matches: 274  
 Percent Similarity: 91.67% Conservative: 12  
 Best Local Similarity: 87.82% Mismatches: 13  
 Query Match: 44.46% Indels: 13  
 DB: 5 Gaps: 2  
 US-10-631-958-10 (1-537) x BQ057191 (1-1047)

Qy 133 LeuLeuValPheIleAsnProPheGlyGlyGlyGlnGlyLysArgIleTyrGluArg 152  
 Db 13 TTACTGTGATTTATCAACCCGTTTGGAGGAAAGGCAAGGCAAGCGGATATATGAAGA 72  
 Qy 153 LysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHis 172  
 Db 73 AAAGTGGCACCACTGTTTCACCTTAGCCTCCATCCACCTGACATCATCGTTACTGAACAT 132  
 Qy 173 AlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleVal 192  
 Db 133 GCTAATCAGGCCCAAGGAGACTCTGTATGACATTAAACATAGACAAATACGACGCGCATCGTC 192  
 Qy 193 CysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGln 212  
 Db 193 TGTGTCGGCGGAGATGGTATGTTTTCAGCGAGGTCTGCACGGTCTGATTTGGGAGACGACG 252  
 Qy 213 ArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArg 232  
 Db 253 AGAGAGCCCGGGTTCAGCACAGAACCCCGGGGTGTGTGGTCCCGACTAGCTCCGG 312  
 Qy 233 IleGlyIleIleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSer 252  
 Db 313 ATTTGAATCATTTCCCGCAGGGTCAACGAGACTGGGTGTGTACTTCCACCGTGGGCACGACG 372  
 Qy 253 AspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSer 272  
 Db 373 GACGACAGAAACCTCGGCGCTGCATATGTTGTTGGGACTCGCTGGCATGGATGTGTCC 432  
 Qy 273 SerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPhe 292  
 Db 433 TCAGTCCACCAACACAGCACACTCTTCGCTACTCCGCTGCTCCCTGTGGGTACGGGTTC 492  
 Qy 293 TyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAsp 312  
 Db 493 TACGGGGACATCATCAAGGACAGTGAAGAAACGGTGGTGTGGTCTTGGCAGATACGAC 552  
 Qy 313 PheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLe 332  
 Db 553 TTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGAGGACAGTGTCTTCCT 612  
 Qy 332 uProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCys 351  
 Db 613 CCTGACACACACACCGTGGGATCTCCCAAGGATAGGAACCCCTGCCCGGACGATGC 672  
 Qy 352 PheValCysArgGlnSerLysGlnLeuGluGluGluGlnLysLysAlaLeuTyr 370  
 Db 673 TTTGTTTTCGGCGCAAGCAAGCAGCAGTGGAGGAGGAGGCCCAAGAAACCACTTGTAT 732  
 Qy 371 GlyLeuGluAlaAlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAl 390  
 Db 733 GGTTTGAACCTCGGGAAGACGGGACAGGATGGGCAATCTTCTGGGGGGAGATTCTTGGC 792  
 Qy 390 aIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSer---P 409  
 Db 793 CATCCATGCCCAACACCTGCTGCTGTGTCGCGGACCCCGGAGGCGCTTCTCCCC 852  
 Qy 409 roAlaAlaHisLeuGlyAsp---GlySerSerAspLeuIleLeuIleArgLysCys-S 427

Db 853 CGGCTGCCCAATTGGGAGAGGGTCTCTGAACTCTCTCTCAGGAAATGGCT 912  
 QY 427 erArgPheAnPhe 431  
 Db 913 CCCGGTTCATTTT 926

RESULT 8  
 CK000525  
 LOCUS  
 DEFINITION AGENCOURT 16369000 NIH MGC 221 Homo sapiens cDNA clone EST 26-NOV-2003  
 IMAGE:30708597 5', mRNA sequence.

ACCESSION CK000525  
 VERSION  
 KEYWORDS  
 SOURCE CK000525.1 GI:38526559  
 EST.

ORGANISM Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 713)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 Tissue Procurement: James Martin, University of Iowa  
 cDNA Library Preparation: M. Bento Soares, University of Iowa  
 Tissue Procurement: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: NDAM1075 row: a column: 22  
 High quality sequence stop: 689.

## FEATURES

source  
 location/Qualifiers

1..713  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30708597"  
 /lab\_host="DH10B Tona"  
 /clone\_lib="NIH MGC 221"  
 /note="Organ: mixed; Vector: pYX-Asc; Site 1: EcoRI;  
 Site 2: NotI; Library is oligo-dT primed and directionally  
 cloned. Denatured RNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with oligo-dT  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to mRNA size fraction, ligated with  
 EcoR I adaptor, digested with Not I and then cloned  
 directionally into pYX-Asc vector. Average insert size  
 4-5Kb. Adaptors 5' (AATCGGCACGAGG)3' and 5'd  
 (CTCGTCCG)3'. 3' linker sequence - CGGCGCTGAGACC T18.  
 Sequencing primers 3' end: T3 promoter primer 5'd  
 (ATTACCTCCTCACTAAGGGA)3'. 5' End: T7 promoter primer 5'd  
 (TATACACTCACTATAGG)3'. Library was constructed in the  
 laboratory of M. Bento Soares. Note: this is a NIH\_MGC  
 Library"

## ORIGIN

Alignment Scores:  
 Pred. No.: 4,426-116 Length: 713  
 Score: 1275.00 Matches: 235  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 44.15% Indels: 0  
 DB: 7 Gaps: 0

US-10-631-958-10 (1-537) x CK000525 (1-713)

QY 277 AnSerThrLeuLeuArgTy-SerValSerLeuLeuGlyTyArgPheTyGlyAspIle 296

Db 9 AACAGCACATCCCTTCGCTACTCCGTGCTCCCTGGGCTACGGCTTCTACGGGACATC 68  
 QY 297 IleYsAspSerGluLeuYsLeuArgTrpLeuGlyLeuAlaArgTyAspPheSerGlyLeu 316  
 Db 69 ATCAAGGACAGTGAGAGAAACGGTGGTTGGTCTTGGCAGATACGACTTTTCAGGTTA 128  
 QY 317 LysThrPheLeuSerHisHisCysTyTrpGluGlyThrValSerPheLeuProAlaGlnHis 336  
 Db 129 AAGACCTTCCTCCACCACTGCTATGAAGGACAGTGTCTTCTCCCTCCGTCACAACAC 188  
 QY 337 ThrValGlySerProAtqAspArgLysProCysArgAlaGlyCysPheValCysArgGln 356  
 Db 189 ACGGTGGGATCTCAAGGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTTGCAGGCAA 248  
 QY 357 SerYsGlnGlnLeuGluGluGlnLysYsAlaLeuTyGlyLeuGluAlaAlaGlu 376  
 Db 249 AGCAAGCAGCAGCTGGAGGAGGAGCAGAAAGACACTGTATGGTTTGGAGCTCGGAG 308  
 QY 377 AspValGluGluTrpGlnValValCysGlyYsPheLeuAlaIleAsnAlaThrAsnMet 396  
 Db 309 GACGTGGAGGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAACATG 368  
 QY 397 SerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGly 416  
 Db 369 TCCTGTGCTTGTCCGGGAGCCCGCAGGGGCTCTCCCGGCTGCCACTTGGAGACGGG 428  
 QY 417 SerSerAspLeuIleLeuIleArgLysCysSerArgPheAnPheLeuArgPheLeuIle 436  
 Db 429 TCTTCTGACCTCATCTCCATCCGGAATGCTCCAGGTTCATTTCTCAGATTCTCATC 488  
 QY 437 ArgHisThrAnGlnGlnAspGlnPheAspPheThrPheValGluValTyArgVallys 456  
 Db 489 AGGCACACCAACACGACGACGACAGTTTGCATTCATCTTTTGTGAAGTTTATCGCGTCAAG 548  
 QY 457 LysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLeuGlyGly 476  
 Db 549 AAATTCAGTTTACGTGAAGCACATGGAGGATGAGGACGACGACCTCAAGAGGGGGG 608  
 QY 477 LysYsArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSerAsn 496  
 Db 609 AAGAAGCGCTTTGGGCACATTTGGCAGCAGCCACCCCTCTCTGCTGTCACCGTCTCCAC 668  
 QY 497 SerSerTrpAnCysAspGlyGluValLeuHisSerProAlaIle 511  
 Db 669 AGCTCCTGGAACCTGCAGCGGAGGTCTCTGCACAGTCTCTGCATC 713

## RESULT 9

CD655311

LOCUS

DEFINITION

AGENCOURT 14552675 KIA Human H1 Embryonic Stem Cell cDNA Library

(Long) Homo sapiens cDNA clone IMAGE:30426593 5', mRNA sequence.

ACCESSION

CD655311

VERSION

CD655311.1 GI:31895467

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 820)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Irene Gulin

CDNA Library Preparation: Yulan Piao and Mahendra Rao, NIA

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC c

lone distribution information

can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>  
 Plate: NDAMS12 row: k column: 18  
 High quality sequence stop: 673.

## FEATURES

source

1. 820  
 /location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30426593"  
 /tissue\_type="Embryonic Stem cells"  
 /cell\_line="WA01"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"

/note="vector: pCMV-Sport6; Site\_1: NotI; Site\_2: SalI;  
 This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01 cell line. Undifferentiated human ES cell line WA01/H1 was obtained from WiCell Research Institute, Inc., Madison, WI, cultured according to their instructions, on MEF feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UTF, TERT, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1, TUBB3, NES, GPAP, and EOMES. When confluent (18-10 days after plating), the ES cells from 4 X 6cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with Trizol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID:11544199]) Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTCTTAGATCGGAGCGCGCCCTTTTCTTTT-3'] from 3.4g of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker L1-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by extraction-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S for 25 cycles. The products were purified by phenol/chloroform extraction and Centricion-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The average insert size is about 3.6kb."

## ORIGIN

## Alignment Scores:

Pred. No.: 1.71e-114 Length: 820  
 Score: 1260.00 Matches: 250  
 Percent Similarity: 93.12% Conservative: 7  
 Best Local Similarity: 90.58% Mismatches: 12  
 Query Match: 43.63% Indels: 8  
 DB: 6 Gaps: 5

US-10-631-958-10 (1-537) x CD655311 (1-820)

Qy 247 SerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSer 266  
 Db 12 TCCACCGTGGCAGCAGCAGCAAAACCTCGCGCTGCATATCGTTGTTGGGACTCG 71  
 Qy 267 LeuAlaMetAspValSerValHisAsnSerThrLeuLeuAgtTyrSerValSer 286  
 Db 72 CTGGCCATGGATGTCTCAGTCACCACCAACAGCACACCTCTCGCTACTCCGTGCC 131  
 Qy 287 LeuLeuGlyTyrGlyPheTyrGlyAspIleLeuLysAspSerGluLysArgTrpLeu 306  
 Db 132 CTGCTGGGCTACGGCTTCTACGGGGACATCATCAAGGACGTGAGAAACGGTGGTGG 191  
 Qy 307 GlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGlu 326  
 Db 192 GGTCTTGCCAGATAGCACTTTTCAGGTTTAGAGACCTTCTCTCCACCACCTGCTATGAA 251

Qy 327 GlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysPro 346  
 Db 252 GGCACAGTGTCTCTCTCCCTGCACACACACCGGTGGATCTCCAGGGATAGGAAGCCC 311  
 Qy 347 CysArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLeuGluGluGlnLys 366  
 Db 312 TCCCGGGCAGGATGCTTTGTTTGCAGGCAAAAGCAGCAGCTGGAGGAGGAGCAAG 371  
 Qy 367 LysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlnTrpGlnValValCysGly 386  
 Db 372 AAAGCACTGTATGTTTGGAAAGCTGCGAGAGCGTGAGGAGTGCGAAGTCGTCTGTGGG 431  
 Qy 387 LysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGly 406  
 Db 432 AAGTTTCTGGCCATCAATGCCACAAACATGCTCTGTGCTTGTGCGCGAGCCCCAGGGGC 491  
 Qy 407 LeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuLeuLeuLeuLysCys 426  
 Db 492 CTCTCCCGGCTGCCACTTGGGAGACGGTCTTCTGACCTCATCTCATCCGGAATGC 551  
 Qy 427 SerArgPheAsnPheLeuArgPheLeuLeuArgHisThrAsnGlnGlnAspGlnPheAsp 446  
 Db 552 TCCAGGTTCAATTTCCTGAGATTTCCTCATCAGGCACACCAACAGCAGGACCATGTTTGAC 611  
 Qy 447 PheThrPheValGluValTyrArgValLysLysPheGlnPheThrSerLysHisMetGlu 466  
 Db 612 TTCATTCTTGAAGTTTATCGGTCAAGAAATCCAGTTTAGTCTGAAGACACATGGAG 671  
 Qy 467 AspGluAspSerAspLeuLysGluGlyGlyLysLysArgPheGlyHisIleCysSerSer 486  
 Db 672 GATCAGCAGCAGCACCTCAAGGAGGGGGAAGAGCGCTTGGGCACAT---TGCAGCAGC 728  
 Qy 487 HisProSerCysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeu 506  
 Db 729 ---ACCTCTGCTGCTGCAGC---TCTCAGAGCTCTG---GACTCGCAGCGGAGTCTG--- 777  
 Qy 507 HisSerProAla---IleGluValArgValHisCysGlnLeuValArg 521  
 Db 778 -----CCAGCTGCATCAGTCAAGTCATGCATGCATGTCGATCGCCAGA 819

CF135528 732 bp mRNA linear EST 09-SEP-2003  
 UT-HF-BNO-amf-g-10-0-UI.r1 NIH\_MGC\_50 Homo sapiens cDNA clone  
 IMAGE:3050211 5', mRNA sequence.  
 CF135528  
 CF135528  
 EST.  
 CF135528.1 GI:33250972  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 732)  
 Ronaldo,M.F., Lennon,G. and Soares,M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genome Res. 6 (9), 791-806 (1996)  
 97044477  
 8889548  
 Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 Tissue Procurement: Louis Staudt  
 cDNA Library prepared by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Distribution information can be found at  
<http://genome.uiowa.edu/distribution/humanfl.html>

JOURNAL  
 MEDLINE  
 PUBMED  
 COMMENT



```

FEATURES
  source
    Seq primer: pYX-5
    Location/Qualifiers
      1..732
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:3090211"
        /tissue_type="lymph"
        /cell_type="germinal center B cells"
        /cell_line="MGC85"
        /lab_host="DH10B (lTI)"
        /clone_lib="NIH MGC 50"
        /note="Vector: pYX-5; Site 1: NotI; Site 2: Eco RI;
        Constructed from size fractionated cytoplasmic mRNA
        (3.5-4.4kb). Directionally cloned. Cells provided by
        Louis M. Staudt, Ph.D. Library preparation by Maria de
        Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."
      ORIGIN
        Alignment Scores:
          Pred. No.: 5.62e-113 Length: 732
          Score: 1244.00 Matches: 231
          Percent Similarity: 99.57% Conservativeness: 0
          Best Local Similarity: 99.57% Mismatches: 0
          Query Match: 43.07% Indels: 1
          Db: 7 Gaps: 0
        US-10-631-958-10 (1-537) x CF135528 (1-732)
      QY 307 GlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisCysTyrGlu 326
      Db 9 GGTCTTGCAGATACAGCTTTTCAGGTTTAAAGACCTTCCTCCACCACCTGCTATGAA 68
      QY 327 GlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysPro 346
      Db 69 GGGACAGTGTCTCTCTCCCTGCACACACACGGTGGGATCTCCAAAGGGATAGGAAGCC 128
      QY 347 CysArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLeuGluGluGlnLys 366
      Db 129 TGCCGGCAGGATGCTTTGTTGCGAGCAAGCAAGCAGCAGCTGGAGGAGGAGGAGAAG 188
      QY 367 LysAlaLeuTyrGlyLeuGluAlaGluAspValGluGluTyrGlnValValCysGly 386
      Db 189 AAAGCAGCTATGTTGTTGAAGCTCGGAGGACGTTGGAGGATGGCAGTCTGCTGTGG 248
      QY 387 LysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGly 406
      Db 249 AGTTTCTGGCCATCAATCCACCAACATGCTCTGTGCTTGTGCGCGAGCCCAAGGGGC 308
      QY 407 LeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuLeuLeuLeuArgLysCys 426
      Db 309 CTCTCCCGGCTGCCACCTTGGGAGACGGGTCTTCTGACCTCATCTCATCGGAATGC 368
      QY 427 SerArgPheAsnPheLeuArgPheLeuLeuArgHisThrAsnGlnGlnAspGlnPheAsp 446
      Db 369 TCCAGGTTCAATTTCTGAGATTTCTCATCAGGCACACCAACAGCAGACCAAGTTTGCAC 428
      QY 447 PheThrPheValGluValTyrArgValLysLysPheGlnPheThrSerLysHisMetGlu 466
      Db 429 TTCACCTTTGTTGAAGTTTATCGGTCAAGAAATTCAGTTTACGTCGAGGACATGAG 488
      QY 467 AspGluAspSerAspLeuLys-GluGlyLysLysArgPheGlyHisIleCysSerSe 486
      Db 489 GATGAGGACAGGACCTCAAGGGAGGGGGAAGAGCGCTTTGGGCACATTTGCGAGCAG 548
      QY 486 rHisProSerCysCysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluValLe 506
      Db 549 CCACCCCTCTGCTGTCAGCAGCTTCCCAACAGCTCTTGGAACTCGCAGCGGGAGTCTCT 608
      QY 506 rHisSerProAlaIleGluValArgValHisCysGlnLeuValArgLeuPheAlaArgG 526
      Db 609 GCACAGCCCTGTCATCGAGGTGAGATTCACCTGCCAGCTGTGTTGCACTCTTTGCAAGG 668
  
```

QY 526 ylleclugluhanProLysProAspSerHisSer 537  
 Db 669 GATTGAAGAGAAATCCGAAGCCAGACTCACAGC 702

RESULT 11  
 LOCUS CB246749  
 DEFINITION UI-M-F10-cdx-b-10-0-UI.r1 NIH\_BMAP\_F10 Mus musculus cDNA clone  
 ACCESSION IMAGE:6835595 5', mRNA sequence.  
 VERSION CB246749  
 KEYWORDS CB246749.1 GI:28368393  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 758)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs@mail.nih.gov  
 Tissue Procurement: Dr. Jim Lin, University of Iowa  
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 This clone was contributed by the Brain Molecular Anatomy Project  
 (BMAP)

Seq primer: pYX-5.  
 Location/Qualifiers  
 1..758  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:6835595"  
 /tissue\_type="whole brain"  
 /dev\_stage="embryo 12.5dpc"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_lib="NIH\_BMAP\_F10"  
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: Eco I;  
 Site 2: Not I; The library was constructed according  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured RNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with oligo-dT  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to mRNA size fraction, ligated  
 with EcoR I adaptor, digested with NotI and then cloned  
 directionally into pYX-Asc vector. The library tag  
 sequence located between the Not I site and the polyA tail  
 is CAGCCACGAC. This library was created for the University  
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the  
 Developing Mouse Nervous System', supported by National  
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,  
 program coordinator."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 4.54e-109 Length: 758  
 Score: 1205.00 Matches: 223  
 Percent Similarity: 95.63% Conservativeness: 18  
 Best Local Similarity: 88.49% Mismatches: 11  
 Query Match: 41.72% Indels: 0  
 Db: 6 Gaps: 0  
 US-10-631-958-10 (1-537) x CB246749 (1-758)

QY 198 GlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyVal 217  
 Db 1 GGCATGTTTCAGCGAGGTGCTGCATGGGGTATGGGAGGACGACGAGCGCGCTGGTATC 60



```
Qy 218 AspGlnAenHisProArgAlaValLeuValProSerSerLeuArgLleGlyVilePro 237
|||
Db 61 GACCCCATACCCCGGAGCGTGTGGTCCAGTACCTCAGATCGGATCATACCC 120
|||
Qy 238 AlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSer 257
|||
Db 121 GCAGGGTCCACAGATTGTGTGTACTCAACAGTGGGCACAAACGACGACAGACATCG 180
|||
Qy 258 AlaLeuHisLleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsn 277
|||
Db 181 GCATTTCACATCATATTATGGGAGCTCACTGGCAATAGACGTGTCTCTGTGCATACCAT 240
|||
Qy 278 SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspLleLle 297
|||
Db 241 AACACGCTGCTGGGTACTCGGTTCTCTGTGGGTACGGTTTCTACGGGGATTTATC 300
|||
Qy 298 LysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys 317
|||
Db 301 AAGGACAGTGAAGAAAGAAAGCGTGGTGGCTCGTCCGGTATGATTTCTCAGGGTTGAAG 360
|||
Qy 318 ThrPheLeuSerHisHisCysTyrGlyThrValSerPheLeuProAlaGlnHisThr 337
|||
Db 361 ACCTTTCTCTCATCATGACTATGAGGGACACTGTCTCTTCTCCAGCACAGCACAG 420
|||
Qy 338 ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer 357
|||
Db 421 GTGGATCTCCACGGGACAAATAACCTTCGCGGCTGGGTCTCGTGTGCGAGCGAGC 480
|||
Qy 358 LysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp 377
|||
Db 481 AAGCAACAGCTGGAAGAAAGAAAGAAAGCCCTGTATGCCCTGGAGAACGCCGAGGAA 540
|||
Qy 378 ValGluGluTrpGlnValValCysGlyLysPheLeuAlaLleAsnAlaThrAsnMetSer 397
|||
Db 541 ATGGAAGAGTGGCAAGTACATGTGGGAAGTTCCTGGCCATCAATGCCACCAACATGTCC 600
|||
Qy 398 CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer 417
|||
Db 601 TGTGCTGTCTCGGAGCCCTGGGGCTGTGTCNCCATTTGCCCATCTGGGAGATGGGTCT 660
|||
Qy 418 SerAspLeuLleLeuLleArgLysCysSerArgPheAsnPheLeuArgPheLeuArg 437
|||
Db 661 TCTGACCTCATCTTATCCGGAAGTCTCCAGGTTCCANCTTCTCTGAGATTCTCATCCGG 720
|||
Qy 438 HisThrAsnGlnGlnAspGlnPheAspPheThrPhe 449
|||
Db 721 CACAGAACCGAGGAGGACGATTCGACTCTACTTTC 756
|||
```

```
RESULT 12
BX952302          653 bp  mRNA      linear      EST 01-MAR-2004
LOCUS            DKEZp781L11183_r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
DEFINITION       DKEZp781L11183_5', mRNA sequence.
ACCESSION        BX952302
VERSION          BX952302.1 GI:43428907
KEYWORDS          EST.
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens
REFERENCE         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS           Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
                  Fobo,G., Han,M. and Wiemann,S.
                  1 (bases 1 to 653)
TITLE             EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., et al.)
JOURNAL           Unpublished (2003)
COMMENT           Contact: MIPS
                  MIPS
```

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;  
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing

consortium of the German Genome Project.

No s1 sequence available.

This clone (DKFZp781L11183) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

#### FEATURES

source

Location/Qualifiers  
1..653  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp781L11183"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="781 (synonym: hlcc4)"  
/notes="Vector: pSport1\_Sfi; Site\_1: SfiIA; Site\_2: SfiIB;  
cDNA-collection"

#### ORIGIN

```
Alignment Scores:      8.84e-107      Length:      653
Pred. No.:            1181.00      Matches:      217
Score:                100.00%      Conservative: 0
Percent Similarity:   100.00%      Mismatches: 0
Best Local Similarity: 100.00%      Indels:      0
Query Match:         40.89%      Gaps:        0
DB:
US-10-631-958-10 (1-537) x BX952302 (1-653)

Qy 297 IleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeu 316
Db 2 ATCAAGGACAGTGAAGAAACCGTGGTGTGGCTTCCAGATACGACTTTTTCAGGTTTA 61
|||
Qy 317 LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis 336
|||
Db 62 AAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTTCCCTGCAACAC 121
|||
Qy 337 ThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGln 356
|||
Db 122 ACGGTGGGATCTCCAGGGATAGGAAGCCCTGCCGGCAGGATGCTTTTGTTCAGGCAA 181
|||
Qy 357 SerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGlu 376
|||
Db 182 AGCAAGCAGCAGCTGGAGGAGGAGCAGAGAAAGACACTGTATGTTTGGAGCTGCGAG 241
|||
Qy 377 AspValGluGluTrpGlnValValCysGlyLysPheLeuAlaLleAsnAlaThrAsnMet 396
|||
Db 242 GACGTGGAGAGTGGCAAGTCTCTGTGGGAGTTTCTGGCCATCAATGCCACAAACATG 301
|||
Qy 397 SerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGly 416
|||
Db 302 TCCTGTGCTTGTGCGCGAGGCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGG 361
|||
Qy 417 SerSerAspLeuLleLeuLleArgLysCysSerArgPheAsnPheLeuArgPheLeuLle 436
|||
Db 362 TCTTCTGACCTCATCTCATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATC 421
|||
Qy 437 ArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLys 456
|||
Db 422 AGGCACACCAACACAGCAGGACAGTGTGACTTCTACTTTTGTGAAGTTATTCGGTCAAG 481
|||
Qy 457 LysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGly 476
|||
Db 482 AAATTCAGTTTACGTGGAAGCACATGAGGATGAGGACAGCGACCTCAAGAGGGGGGG 541
|||
Qy 477 LysLysArgPheGlyHisLleCysSerSerHisProSerCysCysThrValSerAsn 496
|||
Db 542 AAGAAGCGCTTGGGACATTTGCAGCAGCCACCCCTCTCTGCTGCTGACCGCTCTCCAAC 601
|||
Qy 497 SerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaLleGluVal 513
|||
Db 602 AGCTCTGGAACTGCGACGGGAGGCTCTTGACAGAGCCCTGCAATCGAGGTC 652
|||
```

RESULT 13

CN296312  
 LOCUS 17000532197874 GRN\_EB Homo sapiens cDNA 5', mRNA linear EST 16-MAY-2004  
 DEFINITION CN296312  
 ACCESSION CN296312  
 VERSION CN296312.1 GI:47312726  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 661)  
 Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,  
 Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,  
 Lebkowski, J. and Stanton, L.W.  
 Transcriptional characterization elucidates signaling networks that  
 control human ES cell growth and differentiation  
 Nat. Biotechnol. 22 (6), 707-716 (2004)  
 Contact: Brandenberger R  
 Regenerative Medicine  
 Geron Corporation  
 230 Constitution Drive, Menlo Park, CA 94025, USA  
 Tel: 650 473 8658  
 Fax: 650 473 7760  
 Email: rbrandenberger@geron.com  
 Insert Length: 661 Std Error: 0.00.  
 FEATURES  
 Location/Qualifiers  
 1..661  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="embryonic stem cells, embryoid bodies  
 derived from H1, H7 and H9 cells"  
 /clone\_lib="GRN\_EB"  
 /note="oligo dt primed, full-length enriched cDNA library  
 from embryoid body outgrowths derived from hES cell lines  
 H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free  
 conditions."

# ORIGIN

Alignment Scores:  
 Pred. No.: 1,12e-105 Length: 661  
 Score: 1170.00 Matches: 219  
 Percent Similarity: 99.55% Conservative: 0  
 Best Local Similarity: 99.55% Mismatches: 0  
 Query Match: 40.51% Indels: 1  
 DB: 7 Gaps: 0

US-10-631-958-10 (1-537) x CN296312 (1-661)

Qy 278 SerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleile 297  
 Db 3 AGCACACTCTCGCTACTCGGTGTCCTGCTGGGCTACGGCTTCTACGGGACATCATC 62  
 Qy 298 LysAspSerGluLeuLeuArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLys 317  
 Db 63 AAGGACGTGAGAGAAACGGTGGTGGTCTTGCCAGATACGACTTTCAGGTTTAAAG 122  
 Qy 318 ThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThr 337  
 Db 123 ACCTTCTCTCCACCATCTGTATGAGGGACAGTGTCTTCTCTCCCTGCACACACACAG 182  
 Qy 338 ValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSer 357  
 Db 183 GTGGGATCTCCAAGGATAGGAAGCCCTCCCGGCGAGGATGCTTTGTTTTCAGGCAAGC 242  
 Qy 358 LysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAsp 377  
 Db 243 AAGCAGCAGCTGCGAGGAGCAGAGAAAGACACTGTATGGTTTGGAAAGCTCGCGAGGAC 302  
 Qy 378 ValGluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSer 397  
 Db 303 GTGGAGGAGTGCAAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCACCAACATGTCC 362

Qy 398 CysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySer 417  
 Db 363 TGTCTTGTGCGCCGAGCCGCCAGGGGCTCTCCCGGGCTGCCCACTTTGGGAGAGCGGTCT 422  
 Qy 418 SerAspLeuLeuLeuLeuArgLysCysSerArgPheAsnPheLeuArgPheLeuLeuArg 437  
 Db 423 TCTGACCTCATCTCATCCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGG 482  
 Qy 438 HisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLys 457  
 Db 483 CACACCAACAGCAGGACCACTTTGACTTCACTTTTGTGTAAGTTTATTCGGCTCAAGAAA 542  
 Qy 458 PheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGlu-GlyGlyLys 477  
 Db 543 TTCCAGTTTACGTGGAAGCACCATGGAGATGAGACACGACCTCAAGAGGGGGGGAA 602  
 Qy 477 sLysArgPheGlyHisIleCysSerSerHisProSerCysCysCysThrValSerAsn 496  
 Db 603 GAAGCGCTTTGGGCACATTTGCAGACGCCCTCTCTGCTGCTGCACCGTCTCCAC 660

## RESULT 14

CK603033  
 LOCUS CK603033  
 DEFINITION AGENCOURT 17899852 NIH\_MGC\_234 Rattus norvegicus cDNA clone  
 IMAGE:7193195 5', mRNA sequence.  
 ACCESSION CK603033  
 VERSION CK603033.1 GI:41116352  
 KEYWORDS EST.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 812)  
 NTH-MGC http://mgi.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: csapbs-remail.nih.gov  
 Tissue Procurement: Howard Jacobs  
 cDNA Library Preparation: Express Genomics  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LIAM15055 row: b column: 09  
 High quality sequence stop: 657.

## FEATURES

Location/Qualifiers  
 1..812  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10116"  
 /clone="IMAGE:7193195"  
 /tissue\_type="heart, pooled"  
 /lab\_hosts="DH10B Tona"  
 /clone\_lib="NIH\_MGC\_234"  
 /note="Organ: heart; Vector: pExpress-1; Site:1: EcoRV;  
 Site:2: NotI; RNA obtained from pooled heart tissue from a  
 mix of male and female animals at 8 wk old. Tissues were  
 snap-frozen and kept at -80C for two days before RNA  
 extraction and purification (Tri-reagent method). cDNA was  
 primed using oligo-dT primer:  
 5'-pGATAGTTCATGACGACGCGCCGCCCTTC-3'  
 The EcoRV/NotI sites of pExpress-1. Size-selection >1.4kb  
 resulted in an average insert size of 2.2 kb. This primary  
 library is normalized (non-normalized primary library is  
 NIH\_MGC\_233) and was constructed by Express Genomics  
 (Frederick, MD). Note: this is a NIH\_MGC library."

## ORIGIN

```
Alignment Scores:
Pred. No.:      6.14e-94      Length:      812
Score:          1053.50      Matches:    206
Percent Similarity: 85.23%      Conservative: 19
Best Local Similarity: 78.03%      Mismatches: 35
Query Match:    36.48%      Indels:    5
DB:              7          Gaps:      3

US-10-631-958-10 (1-537) x CK603033 (1-812)

Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnGlnArg 20
   |||
Db 15 ATGGCGCAATGGGGCGGCGAGCCGCTGCAATCGGTCTGTGGTGAAGCACAACGC 74

Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40
   |||
Db 75 TGTCCGCTCAGCTTGGAGCCGCGCGGCTCTGCTAGCGTGGTGGCGAGCCCGAGGCC 134

Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal 60
   |||
Db 135 GGGCCCTCGGCCCGCGATCCGATCGCTACTCGGTACAGTGTGAGATCATCGCCGTT 194

Qy 61 GluGluThrAspValHisGlyLysHisGlnGlnSerGlyLysTrpGlnLysMetGluLys 80
   |||
Db 195 GAGGAAAAGAGTCCACCAGAAACAGTCTCCCAATGGCCGATGGCATAGATGGAATAAT 254

Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAla 100
   |||
Db 255 CCGTTTGCATTTACAGATCCACTGTGTGAAGCGAGCTCGACACCCCGCTGGAAGTGGGCA 314

Qy 101 GlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120
   |||
Db 315 CGGTGTACCTTCTGGAGCCCGCAGCAGCAGTGTCTCACTGTGCTGCGACCCCTCGA 374

Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
   |||
Db 375 GAGCTGTGGAGAACCTGACTTCAAGACCGAAGCATTTGCTGTGTTATTTATCAACCCCTTC 434

Qy 141 GlyLysLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
   |||
Db 435 GGAGGAAAGGCCCGAGGCGAAGCACAATCTATGAAAAGAAAGTGGACCTTTGTTCCACTTG 494

Qy 161 AlaSerIleThrThrAspIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180
   |||
Db 495 GCTTCCATCACCATGAGATCATCATTTACTGAGCATGCCAATCAAGCCAGGAGACTTTA 554

Qy 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200
   |||
Db 555 TACGAGATCAACACAGACAGCTATGACGGCATCTGTGTGTGTTGGTGGGCGGCATGTTTC 614

Qy 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220
   |||
Db 615 AGCGAGGTGCTGCGACGGGGTAATTGGGAAGACGACGAGCGCTGTTGTTGACCCCAAT 674

Qy 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240
   |||
Db 675 CACCCCGAGCCGTGTGTCGTCAGTACCCT--AGATCGGCATCATCTCCGCGAGGTAC 732

Qy 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
   |||
Db 733 AGA---TGTGTGTGC---TCTCACCGTGGCACAACG-----AGCAGAACGTCGCTTACC 780

Qy 261 IleValValGly 264
   |||
Db 781 ATCATATTGGGA 792

RESULT 15
BP224560
LOCUS
DEFINITION BP224560 Sugano cDNA library, lymphocyte Daudi Homo sapiens cDNA
clone DAT01067, mRNA sequence.
ACCESSION BP224560
```

```
VERSION BP224560.1 GI:52097465
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 581)
Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
FEATURES
source
1..581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DAT01067"
/cell_type="lymphocyte"
/cell_line="Daudi"
/clone_lib="Sugano cDNA library, lymphocyte Daudi"
/note="Burkitt's lymphoma"
ORIGIN
Alignment Scores:
Pred. No.:      4.96e-90      Length:      581
Score:          1012.00      Matches:    192
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.48%      Mismatches: 0
Query Match:    35.04%      Indels:    0
DB:              5          Gaps:      0

US-10-631-958-10 (1-537) x BP224560 (1-581)

Qy 184 AsnIleAspLysTyrAspGlyIleValCysValGlyAspGlyMetPheSerGluVal 203
   |||
Db 3 AACATAGACAAATACACGCGCATCGTCTGTGTGCGCGAGATGATGTTCAGCGAGGTG 62

Qy 204 LeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArg 223
   |||
Db 63 CTGCACGGTCTGATTGGGAGGACGACAGAGGCGCGGGCTCGACCGAGGCCACCCCGG 122

Qy 224 AlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySerThrAspCys 243
   |||
Db 123 GCTGTGTGTCGCCAGTAGCTCCGGATTGGAATCATTCCTCCGCGAGGTCAACGAGACTGC 182

Qy 244 ValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValVal 263
   |||
Db 183 GTGTGTTTACTCCACCGTGGGCACGACGACGAGAAACCTCGCGGCTGCATATCGTTGTT 242

Qy 264 GlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeuLeuArgTyr 283
   |||
Db 243 GGGGACTCGTGGCCATGGATGTCTCTCAGTCCACACACAGCACACTCTTCGTGTAC 302

Qy 284 SerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGluLysLys 303
   |||
Db 303 TCCGTGTCCCTGCTGGGCTACGGCTTCTAGCGGGACATCATCAAGGACAGTGAAGAGAAA 362

Qy 304 ArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHis 323
   |||
Db 363 CGGTGGTGGGTCTTGCAGATAGACTTTTTCAGGTTTAAAGACTTCTCTCTCCACCAC 422

Qy 324 CysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAsp 343
   |||
Db 423 TGCTATGAAGGGACAGTGTCTCTCTCTCCCTGCACACACACGCGTGGGATCTCCAAAGGAT 482

Qy 344 ArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnLeuGluGlu 363
   |||
```

Db 483 AGGAAGCCCTGCCGGCAGGATGCTTTGTTCAGGCAAGCAAGCAGCTGGAGGAG 542

Qy 364 GluGlnLysLysAlaLeuTyGlyLeuGluAlaGlu 376

Db 543 GAGCAGAAAGAAAGCACTGTATGGTTTGGAAAGCTGCGGAG 581

Search completed: September 6, 2005, 20:29:36  
Job time : 5188.87 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 04:10:49 ; Search time 23.7411 Seconds  
(without alignments)  
2176.332 Million cell updates/sec

Title: US-10-631-958-10  
Perfect score: 2888  
Sequence: 1 MGATGAEPLOSVLVWQQR.....QLVRLFARGIENPKPDSSH 537

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	567.5	19.7	549	2 T33517	hypothetical prote
2	417.5	14.5	1240	2 T05162	hypothetical prote
3	291.5	10.1	458	2 T38776	hypothetical prote
4	248.5	8.6	687	2 S51398	hypothetical prote
5	245.5	8.5	473	2 T19707	hypothetical prote
6	210.5	7.3	624	2 S67059	hypothetical prote
7	189.5	6.6	310	2 AG1665	hypothetical prote
8	176	6.1	310	2 A11293	hypothetical prote
9	147	5.1	303	2 F69795	conserved hypotet
10	136.5	4.7	306	2 A11769	conserved hypotet
11	136	4.7	732	2 T16422	hypothetical prote
12	130.5	4.5	309	2 A11528	conserved hypotet
13	129	4.5	295	2 A83894	hypothetical prote
14	126.5	4.4	306	2 A81394	conserved hypotet
15	125.5	4.3	295	2 D83734	hypothetical prote
16	125.5	4.3	309	2 A11171	conserved hypotet
17	122.5	4.2	433	2 S75948	hypothetical prote
18	115.5	4.0	294	2 G95120	conserved hypotet
19	114	3.9	364	2 F84898	hypothetical prote
20	112	3.9	345	2 E69678	involved in polyke
21	111.5	3.9	311	2 C97990	conserved hypotet
22	105	3.6	333	2 F71006	hypothetical prote
23	103.5	3.6	1028	2 A86719	hypothetical prote
24	102.5	3.5	315	2 A89978	conserved hypotet
25	102	3.5	297	2 F69595	multidrug resistan
26	102	3.5	343	2 JC7183	cathepsin Q (EC 3.
27	100	3.5	650	1 JCI450	fibroblast growth
28	99	3.4	1555	2 T18688	hypothetical prote
29	99	3.4	1973	2 G89608	protein B0272.5 [1

RESULT 1

T33517  
hypothetical protein T10B11.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T33517  
R:Minx, P.; Kemp, K.  
submitted to the EMBL Data Library, October 1998  
A:Description: The sequence of C. elegans cosmid T10B11.  
A:Reference number: Z21363  
A:Accession: T33517  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-549 <MLN>  
A:Cross-references: UNIPROT:Q9TZ11; EMBL:AF098993; PIDN:AAC67466.1; GSPDB:GN00019; CESP:  
A:Experimental source: strain Bristol N2; clone T10B11  
C:Genetics:  
A:Gene: CESP:T10B11.2  
A:Map position: 1  
A:Introns: 26/1; 76/2; 109/3; 159/1; 187/3; 229/3; 353/1; 398/2; 434/1; 467/3

Query Match 19.7%; Score 567.5; DB 2; Length 549;

Best Local Similarity 30.8%; Pred. No. 3.4e-39;

Matches 139; Conservative 91; Mismatches 160; Indels 61; Gaps 13;

Qy	89	VKRARRHRKWAQ--VTFWCPEEOLCHLWLTQLREMLEKLTSPKHLVFNPFQGGKQG	146
Db	123	VYKDKQKWLKQIPVIFYTTSR--DYWHSLLDTTLRRVKNRPANIIIFINPFGNGKA	180
Qy	147	KRIYERKVAFLTLA-SITTDIIIVTEHANAQAKETLYEINIDKY---DGIVCVGGDGMFSE	202
Db	181	QKIFKDNVDFAFFLTPGLRYKVLTERRANHARDYIVEMPPQWSAIDGLSVSGDGLFNE	240
Qy	203	VLHGLIGRTQBSAGVDQNHPRAVLPSSLRIGIIPAGTDCVCYSTVGTSTDAETSAHLIV	262
Db	241	LLSGALLRTQTDGRNIDNPSSHLVTPHIREGIIAGSANSIVSTVHETNDHATSAVHIA	300
Qy	263	VGDSLAMDVSSVHNSLTLLRYSVSLGFGVFFGDIIDKSEKKRWGLARYDSGLKTFLSH	322
Db	301	IGSECNVDVCTVHOHQKLIRISANAISYGVMDVLRDSEYRCLGPIRYQMSALRITIRH	360
Qy	323	HCYEGTSFSPAQHTVGSPRRD-KPCRAGCFCVCRQSKQQLKEEKKALYGLAAEDVSEW	381
Db	361	PIYGMVQFSLSHKENVNPDKQLPCLPCPCVCMK-----PQNDKDYDHWHA----EFT	411
Qy	382	QVWCKGLAINATNWSCACRRSPRGLSPAHLGGSSDLILIRKCSRFNFRFLIRHT--	439
Db	412	HVICCVIPTVTPF-----TPYGLAPFTGIGDGTLDLALVPRISRFRHNNQFMKRVAMY	463
Qy	440	--NQDQDFTFVVEYRVKVFQFTSKHMEDESDLKEGGKKRFCHI CSSHFSFCCCTVSNS	497
Db	464	GGKQLYELD-PSLNCYRVTKWSY-----OPDAQDQDPG-----	495







A;Accession: AG1665  
A>Status: preliminary  
A:Molecule type: DNA  
A;Residues: 1-310 <GLA>  
A;Cross-references: UNIPROT:Q92AQ5; GB:AL592022; PIDN:CAC97095.1; PID:g16414366; GSPDB:G  
A;Experimental source: strain Clip11262  
C;Genetics:  
A;Gene: lin1865

Query Match            6.6%    Score 189.5; DB 2; Length 310;  
Best Local Similarity 21.8%; Pred.No.5.3e-08;  
Matches 91; Conservative 57; Mismatches 139; Indels 131; Gaps 16;

Qy 131 KHLVFTNPFGKGGKKRIYERKVAFLTLAS----ITTDIIIVTEHANAKETLYEINID 186  
||| ||| :  
Db 3 KHARVIYNP-----TSGREHIIKKNLADVLISLEQAGYVTSAHATTAFPGDAKHAAEEAVRD 58  
||| ||| :  
Qy 187 KYDGIVCVGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGSTDCVCY 246  
||| ||| :  
Db 59 RYDLVAAGDGTINEVINGIAEQPYRP-----KVGIPTTGTINDFAIR 101  
||| ||| :  
Qy 247 STVGTSDAETSALHIIVGDSLSANDVSVVHNSTLLRRYSVSLGCGFYGDIKD--SEKKR 304  
:  
Db 102 ALHVPRDIVIKTIILAAQSVMADIGKA--NDT---YFINIGGGRLTELTYDVPSELKT 156  
||| ||| :  
Qy 305 WLGLARYDFSGLTKLFSLHHCEYGVTSFLPAQHTVGSPPDRKPCRCAGFCVCRSQKOLEEE 364  
||| ||| :  
Db 157 MLGOLAYYLXG-----IEMLPs-----LKATKVKVEYD 184  
||| ||| :  
Qy 365 QKXALYGLEAAEDVEEVQVVGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDLILIR 424  
||| ||| :  
Db 185 Q-----GVFEGEVN---FFLLGLTNSIGPEK---IAPDAKLDDGKFSLIIVK 226  
||| ||| :  
Qy 425 KCSRPNFLRF-----TRHTNQDPDTFFVEYRVKFKQFTSKHMEDESDLKEGG 476  
||| ||| :  
Db 227 KVNLAEPIRLVTALRGDHIKEFN-----VIYVKS-EKSVSHSEDK----- 266  
||| ||| :  
Qy 477 KKRFGHICSHSPSCCTVSNSWNCDEGVLAHSPAIEVRVHCQLVRLFARGIEENPKPD 534  
||| ||| :  
Db 267 -----MLINLDGELGGTPEMFRNLKQHIEFFA-SVDDIPATD 303  
||| ||| :

RESULT 8  
Ail1293  
hypothetical protein lmol1753 [imported] - Listeria monocytogenes (strain EGD-e)  
C;Species: Listeria monocytogenes  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C;Accession: Ail1293  
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Feiki, H.  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schluster, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,  
A;Title: Comparative genomics of Listeria species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: Ail1293  
A>Status: preliminary  
A:Molecule type: DNA  
A;Residues: 1-310 <GLA>  
A;Cross-references: UNIPROT:Q8Y6D4; GB:NC\_003210; PIDN:CAC99831.1; PID:g16411207; GSPDB:G  
A;Experimental source: strain EGD-e  
C;Genetics:  
A;Gene: lmol1753

	Qy	187	KYDIGVCGGDMGFSEVLHGILGTQRSGAGVDQHNPRAVLVPSSLRIGIIPAGSTDCVCY	246	
	Dd	59	RFDLVAAAGDGTINEVINGAEEKRP-----KVGLIPTTTDFAR	101	
	Qy	247	STVGTSDAETSALHIVVGDSDLAMDVSSHNNSTLLRYSVSLLYGYFGYGDIINK--SEKKR	304	
	Dd	102	ALHVPRDVIRAKTKIIAAGSQSVAMDIGKA--NET---YFINIGGGRLTETYDVP SRLKT	156	
	Qy	305	WLGARYDFSGLTKTFLSHHCEVTGVSFPLPAOHTVGSPPDRKPACRGFCVRQSQQOLEE	364	
	Dd	157	MLQLAYLLXG-----IEMLPS-----LKATKVKVEYD	184	
	Qy	365	OKKALYLEAEADVEEVQVCGKFLANATNMNSCACRESPRGLSPAHLGDSDDLILR	424	
	Dd	185	Q-----GVFGEVM---FFLLGLTNISGFEK-----IAPDAKLDDKGFSLIIVK	226	
	Qy	425	KCSRFNFRFL 435   :   :   :		
	Dd	227	KWLAEFIRLV 237   :   :   :		
	RESULT 9				
	F69795	conserved hypothetical protein yerQ - Bacillus subtilis			
	C:	Species:	Bacillus subtilis		
	C:	Date:	05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004		
	C:	Accession:	F69795		
	C:	Kunst, F., Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo,			
	C:	A; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter,			
	C:	A; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Fe-			
	Nature	390, 249-256, 1997			
	A:	Authors:	Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi,		
	A:	iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hu,			
	A:	Koetter, P.; Koningsstein, G.; Kroeh, S.; Kumano, M.; Kurita, K.; Lapidus, A.			
	A:	Authors:	Laufer, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.		
	A:	Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.;			
	A:	Rieger, M.; Rivolta, C.; Rocha, E.; Roche, R.; Rose, M.; Sadale, J.; Sato, T.			
	A:	Authors:	Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, Y.; Sekowska,		
	A:	kuechi, M.; Tamakoshi, A.; Tanaka, T.; Terptrata, P.; Tognoni, A.; Tosato, V.			
	T;	Winters, P.; Wipat, A.; Yamamoto, H.; Yanane, K.; Yasumoto, K.; Yata, K.;			
	A:	Authors:	Yoshikawa, H.F.; Zunshein, E.; Yoshikawa, H.; Danchin, A. Bacillus		
	A:	Title:	The complete genome sequence of the Gram-positive bacterium		
	A:	Reference number:	A69580; MUID:98044033; PMID:9384377		
	A:	Accession:	F69795		
	A:	Status:	Preliminary; nucleic acid sequence not shown; translation not shown		
	A:	Molecule type:	DNA		
	A:	Residues:	1-303 <KUN>		
	A:	Cross-references:	UNIPROT_O31502; GB_Z99107; GB_AL009126; NID_G2632866; PIDN		
	A:	Experimental source:	strain 168		
	C:	Genetics:			
	A:	Gene:	yerQ		
	Query Match				
	Best Local Similarity	5.1%; Score 147; DB 2; Length 303;			
	Matches	73; Conservative 54; Mismatches 120; Indels 90; Gaps 15			
	Qy	131	KHLAVFINPFGKGGRRIERYEKVAPL---FTLASITTDIIIVTEHANOAKEYLEYINIDK	187	
	Dd	2	KRARIIYNP----TSGREIFPKHLAQVLQKEQEAGYETSTTHATTCAGDATHAAKEAALRE	57	
	Qy	188	YDGTVCVGGDMGFSEVLHGILGTQRSGAGVDQHNPRAVLVPSSLRIGIIPAGST--DCVCY	246	
	Dd	58	FDLIAAGDGTINEVNGL-----APLD-NRP-----TLGPVPGTTNDFAARA	100	
	Qy	247	STVGTSDAETSALHIVVGDSDLAMDVSSHNNSTLLRYSVSLLYGYFGYGDIINK--SEKKR	304	
	Dd	101	LGIIPREDILKAADVINGVARPIDIQVNG-----QYFINIAGGGRLTETYDVP SKLT	155	
	Qy	305	WLGARYDFSGLTKTFLSHHCEVTGVSFPLPAOHTVGSPPDRKPACRGFCVRQSQQOLEE	364	
	Dd	156	MLQLAYLLXGME-----MLPS-----LRPEVEIEYD	183	
	Qy	365	OKKALYLEAEADVEEVQVCGKFLANATNMNSCACRESPRGLSPAHLGDSDDLILR	424	

```

Query Match      6.1%; Score 176; DB 2; Length 310;
Best Local Similarity 22.2%; Pred.No. 7.1e-07;
Matches 69; Conservative 51; Mismatches 109; Indels 82; Gaps 11;

Qy 131 KHLVLFNPFGGGQGRRIYERVKVAPLFTLAS---ITTDIIVTEHANQAKETLYEINID 186
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 3 KHAVRIYNP---TSGREIIKGNLADVLSILEQAGYVTSAHATTAEPPDDAKAAEEAVRN 58

```





```
Db      162 YKGLEK-----LQI-----SPTDVR----- 178
QY      372 LEAAEDVEEQVVCCKFLAINATNMSCACRRSPRGLGSDLPAAHLGDGSSDLILIRKCSRFNF 431
Db      179 IEYDGKLFEGEIM--MFLVSN-TNSVGGFER---LAPNASLRDGMDFDIIIVKTSPPFP 231
QY      432 LRFL-----IRHTNQDQDFDTFVEYVRK 456
Db      232 LHLAALRGEHIKHPK-----VLYVQANRIK 258
```

Search completed: September 3, 2005, 04:38:44  
Job time : 25.7411 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2005, 03:56:01 ; Search time 29.3937 Seconds  
(without alignments)  
1363.781 Million cell updates/sec

Title: US-10-631-958-10

Perfect score: 2888

Sequence: 1 MGATGAEPLOSVLWVKQOR.....QLVRLFARGIENPKPDSHS 537

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pgp.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pgp.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pgp.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pgp.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pgp.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pgp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	801.5	27.8	687	4	US-09-270-767-45874 Sequence 45874, A
2	585.5	20.3	359	4	US-09-270-767-46720 Sequence 46720, A
3	306	10.6	490	4	US-10-053-510-19 Sequence 19, Appl
4	302.5	10.5	299	4	US-09-270-767-61405 Sequence 61405, A
5	301	10.4	524	4	US-10-053-510-20 Sequence 20, Appl
6	271.5	9.4	618	4	US-09-970-516-4 Sequence 4, Appl
7	271.5	9.4	618	4	US-09-817-676A-14 Sequence 14, Appl
8	267.5	9.3	617	4	US-09-817-676A-12 Sequence 12, Appl
9	266.5	9.2	384	4	US-09-959-897-2 Sequence 2, Appl
10	264.5	9.2	384	4	US-09-970-516-2 Sequence 2, Appl
11	263.5	9.1	384	4	US-09-949-016-7026 Sequence 7026, Ap
12	263.5	9.1	384	4	US-09-796-487-3 Sequence 3, Appl
13	260	9.0	388	4	US-09-817-676A-15 Sequence 15, Appl
14	260	9.0	388	4	US-09-796-487-2 Sequence 2, Appl
15	255.5	8.8	368	4	US-10-053-510-21 Sequence 21, Appl
16	254	8.8	373	4	US-09-796-487-5 Sequence 5, Appl
17	254	8.8	381	4	US-09-796-487-1 Sequence 1, Appl
18	254	8.8	381	4	US-09-796-487-4 Sequence 4, Appl
19	247.5	8.6	382	4	US-09-970-516-6 Sequence 6, Appl
20	244.5	8.5	392	4	US-09-796-487-6 Sequence 6, Appl
21	242	8.4	424	4	US-09-796-487-8 Sequence 8, Appl
22	228	7.9	536	4	US-09-248-796A-15859 Sequence 15859, A
23	206.5	7.2	204	4	US-09-796-487-9 Sequence 9, Appl
24	198	6.9	403	4	US-09-796-487-7 Sequence 7, Appl
25	161.5	5.6	312	4	US-09-949-016-9811 Sequence 9811, Ap
26	153	5.3	313	4	US-09-107-532A-7154 Sequence 7154, Ap
27	148	5.1	119	4	US-09-205-258-788 Sequence 788, App

28	146	5.1	372	4	US-09-603-208A-264 Sequence 264, App
29	144	5.0	319	4	US-09-270-767-38102 Sequence 38102, A
30	144	5.0	319	4	US-09-270-767-53319 Sequence 53319, A
31	126.5	4.4	293	4	US-09-205-258-328 Sequence 328, App
32	123.5	4.3	304	4	US-09-134-000C-5366 Sequence 5366, Ap
33	119.5	4.1	283	4	US-09-634-238-373 Sequence 373, App
34	118	4.1	303	4	US-09-107-532A-5133 Sequence 5133, Ap
35	117	4.1	160	4	US-09-270-767-62339 Sequence 62339, A
36	116.5	4.0	294	4	US-09-583-110-3740 Sequence 3740, Ap
37	116.5	4.0	315	4	US-09-107-433-3121 Sequence 3121, Ap
38	113.5	3.9	345	4	US-09-134-000C-3813 Sequence 3813, Ap
39	110	3.8	324	3	US-09-134-001C-4774 Sequence 4774, Ap
40	107.5	3.7	325	4	US-09-107-532A-6786 Sequence 6786, Ap
41	107	3.7	63	4	US-09-959-897-45 Sequence 45, Appl
42	104.5	3.6	64	4	US-09-959-897-37 Sequence 37, Appl
43	103.5	3.6	316	4	US-09-710-279-1350 Sequence 1350, Ap
44	100	3.5	974	4	US-09-883-134-7 Sequence 7, Appl
45	98.5	3.4	297	4	US-09-134-000C-3616 Sequence 3616, Ap

ALIGNMENTS

RESULT 1  
US-09-270-767-45874  
; Sequence 45874, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270.767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 45874  
; LENGTH: 687  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-45874

Query Match	27.8%	Score	801.5	DB	4	Length	687
Best Local Similarity	31.3%	Pred. No.	1.6e-74				
Matches	197	Conservative	97	Mismatches	206	Indels	129
Gaps	16						
Qy	11	QSVLWVKQORCAVSL	---	PARALLRWRS	PGPGAGAGADACSV	PSVSI	IIAVEETDVHGK 68
Db	59	QQLVWERLQIKQSPQNEAKAPL	----	PPDSPAQPGGICSGPQSHVLHD	--DVVSI	111	
Qy	69	HQSGKWKQMEKP	-----	YAFVHCVKRARR	-----	HRKWAQ	101
Db	112	RSQDTKASSLKPPSPGSRSSGCGDVAQCKPTSQVLTINYAMRLSKSQTDGKRWELRR	171				
Qy	102	VTFCPEQOLCHLWLTREMLEKLT	---	SRPKLLVFINPFGGKGQKRYERKVAFLF	158		
Db	172	LTFFNSDPYIVRQWDQLQRLHSSPTRMKRVRLLVFINPYGKKGAGQYERHVRIF	231				
Qy	159	TLASITTDIIIVTEHANOAKETLYEINIDKYDGVVCVGDGMFSEVLHGLIGRTORSAGVD	218				
Db	232	QLAGVDATCITTORANQVKOILLSHDLGVYDAVCCVGDGTVAEVINGLIIFROMRELGLD	291				
Qy	219	QNHPRAVLVPSLRIGIIPAGSTDCVCTVGTGTSDAETSAHLHVVGDSLAMDVSSVHNS	278				
Db	292	EQRPPYIPRP-ALPVGVIPAGSTDTIAYSMHGTADVTAATHVLGQHRGLDVCVSNGQ	350				
Qy	279	TLLRYSVLLGYGYGDI	IIKDSKKRWLGRLARYDFSGLTPLSHHCHYEVTVSFL	-----	332		
Db	351	SLLRFCAVLSVYGLGDVAAQSENTRYRMWMPRRYISGVKAPLNNRGDAELRMLEEPDLL	410				
Qy	333	---PAQTVGSPRD	-----	RKPCRAGCFVC	-----	RQSKQOLEEBOKKALYGL	373
Db	411	LTTPLEDIPQSDVCSLGSVPVSVVCYANCQRCFASSIQBQRSSLFIOESKEA	-----	E	466		



```
Qy 374 AAEDVE-----EWOVCGKFLAINAT 394
Db 467 RNQVETEDSHLASEAALLPRPGNLRPTGSISSMRNLGNDQWKVYVGNFMCGA 526
Qy 395 NMSCACRRSPGLSPAHLGSSDLILIKCRSFNFLIRHTNQO-QDFDFTFVEVY 453
Db 527 NITCARSPNGISRYSHLGGCLDLILVKKTSLNNVRFLLNTAGRSGDIRNLPFVEVY 586
Qy 454 RVKKFQFTSKHMEDESDLKEGKKRFQHC-----SSHPSCCCTVSNSWNCDE 504
Db 587 RTRFRFTFASSEEDYSLAGS-----COPITPPEMTAHS--STFSSWNCDE 635
Qy 505 VLHSPAIEVRVHCQLVRLFARGIENPKP 533
Db 636 VVTDLITMRSHCQLIEVFMRGPHYSYKP 664

RESULT 2
US-09-270-767-46720
; Sequence 46720, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46720
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46720

Query Match 20.3%; Score 585.5; DB 4; Length 359;
Best Local Similarity 35.5%; Pred. No. 2.4e-52;
Matches 125; Conservative 55; Mismatches 103; Indels 69; Gaps 6;

Qy 129 RPKHLLVINPFGGQGGKRYERKVPAPLFTLASITTDIIIVTEHANQAKETLYEINDKY 188
Db 13 RVRLLVFINPYGGRKAGAQTYERHVRIFQLAGVDATCITTORANQVKDILLSHDLGVY 72
Qy 189 DGIVCGGDGMFSEVLHGLIGRTORSAGVDQNHPRAVLPVPSLSRIGIIPAGSTDCVCYST 248
Db 73 DAVCVGGDGTVAEVINGLI FRQRELGLDQRPPYIPRP-ALPGVIPAGSTDTIAYSM 131
Qy 249 VGTSDAETSAHLIVVGDSLAMDVSGSVHNSLTLLRYSVSLGFGFYGDIIDKSEKKRWLGL 308
Db 132 HGTADVRTAAIHVILGQHRGLDVCVSNQSLRFPASVLSYGVYLGVDVAQAQSNYRWMP 191
Qy 309 ARYDFSGLKTFLSHHCYEGTVSFL-----PAQHTVGSPRD-----RKPCCAGC 351
Db 192 RRYEYSGVKAFNNRNGYDAELRMLDEPDLTLTPLEDIPQSPDVSCLSGESVPSVCYANC 251
Qy 352 FVC-----ROSQQLKEEOKKALYGLEAAEDVE-----ERNOQVETEDSHLASEAALLPRPGNLR 379
Db 252 QRCSPASSIQORSLSLFTQESKEA-----ERNOQVETEDSHLASEAALLPRPGNLR 307
Qy 380 -----EWOVCGKFLAINATNMSCACRRSPRLGSPAHLGPG 416
Db 308 LPTGSISSMRNLGNDQWKVYVGNFPMICANITCARSNGISRYSHLGDG 359

RESULT 3
US-10-053-510-19
; Sequence 19, Application US/10053510
; Patent No. 6830881
; GENERAL INFORMATION:
; APPLICANT: Saba, Julie D.
; APPLICANT: Fyret, Henrik
; TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
; POLYNUCLEOTIDES AND MODULATING AGENTS AND
```

```
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 200116.402C2
; CURRENT APPLICATION NUMBER: US/10/053,510
; CURRENT FILING DATE: 2002-01-17
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-053-510-19

Query Match 10.6%; Score 306; DB 4; Length 490;
Best Local Similarity 24.6%; Pred. No. 6.8e-23;
Matches 119; Conservative 68; Mismatches 165; Indels 132; Gaps 19;

Qy 110 QLCHLWLOTLEMLE-KLTSRP----KHLVFNPFPGKGQGGKRYERKVPAPLFTLASIT 164
Db 12 QTAEMWHHTIRKHKRNGSGSPADCGKQLLILLNPKSGSGKRGELFKQVAPLLTEAEVQ 71
Qy 165 TDIIVTEHANOAKE---TLYEINIDKYDGIVCVGGDGMFSEVLHGLIGRTORSAGVDONH 221
Db 72 YDQITTHPOYAKEFVETRRDL-LTRYSGIVVASGGDLFYEVNLGLMER-----MDW 122
Qy 222 PRAVLVPSLSRIGIIPAGSTDCV-----CYSTVGTSDAETSALHIVVGDSLAMDYSSV 274
Db 123 RRAC---RELPLGIIPCGSGNGLAKSVAHHCNEPEYKPIHLHATLTCMAGKSTPMDVVRV 179
Qy 275 H---HNSTLLRYSVSLGFGFYGDIIDKSEKKRWLGLARYDFGLKTLFLSHHCYEGTVSF 331
Db 180 ELATRDKHFVMSYFLSVGWLIIADIDIESRLRSIGAQRTLWAKRLIGLSYKGRVSY 239
Qy 332 -----LPAQHTVGSPRDRKPCRAGCFVCRQSKQQLKEEOKKALYGLEAA 375
Db 240 LLGKKEKPEPVEAARELPASTAGIRSSLPLNAGEF-----HDLPEEEGEAVLDGEQFA 295
Qy 376 EDV-----EEQVVC-----GKFLAINATNMSCACR-----RSP 404
Db 296 DAISLDRSVYRQHADSWHSAHSRRRTAYYSLGGPSMRNSRMSISQRIEANAABFAERVP 355
Qy 405 RG-----LSPAHLGDSDDLILIRK- 425
Db 356 TGTIPPLQMLLSLDGWCEDGDFVMVHAAAYTHLSSDVFPAFESRLDDGLIYLVIRRG 415
Qy 426 CSRENFELIRHTNQDQF---DFTFVEYVYKFKFOTSKHMEDESD---LKEGKGKR 479
Db 416 VSRHQLNFMIL-NLNAGTHLPIGEDPFIKVPVPCRAFR-----IEPSSSDGILVVDGERVE 469
Qy 480 FQHI 483
Db 470 YGPI 473

RESULT 4
US-09-270-767-61405
; Sequence 61405, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61405
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-61405

Query Match 10.5%; Score 302.5; DB 4; Length 299;
Best Local Similarity 29.2%; Pred. No. 7.3e-23;
```

```
Matches 85; Conservative 35; Mismatches 82; Indels 89; Gaps 9;
QY 317 KTFLLHHCYGVTVFL-----PAQTVGSPRD-----RKPCRACGFCV----- 354
Db 1 KAFLLNRGYDABRLMELEPDLLLTTPLEDIPQSPDSVCSGLGESVPSVCVANCQRCSPASS 60
QY 355 ---RQSKOOLEEOKKALYGLEAAEDVE----- 379
Db 61 IOEQRSSUFIQESKEA-----ERNQOVETEDSHLAASEAALLRPRPRGNLRLPTGSISS 116
QY 380 -----EQVQVCGFLAINATMNSCACRRSPRGLSPAHLGDGSSDLILIRKCSRNFNL 432
Db 117 MENLGNQWKVVRGNFFMCGANITACARSPNGISRYSHLGDGCLDLILYKKTSLNANV 176
QY 433 RPLIRHTNQ--DQDFTFVEVYRVKKFOFTSKHMEDESDLKKEGKKRFGHIC----- 484
Db 177 RPLNTAGRSGDIRLPPFEVYRTREFRFTPSASEEDYSLAGS-----CQPIPTPE 228
QY 485 ---SSHPSCCTVSNSWNCDEGLVHSPAIEVRVHCOLVRLFARGIEENPKP 533
Db 229 EMTAHSS---STEFSSWNCDEGEVTDLOITMRSHCOLIEVFMRGPHYSYK 276

RESULT 5
US-10-053-510-20
; Sequence 20, Application US/10053510
; Patent No. 6830881
; GENERAL INFORMATION:
; APPLICANT: Saba, Julie D.
; TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPTPTIDES,
; TITLE OF INVENTION: POLYNUCLEOTIDES AND MODULATING AGENTS AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 200116.402C2
; CURRENT APPLICATION NUMBER: US/10/053,510
; CURRENT FILING DATE: 2002-01-17
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-053-510-20

Query Match 10.4%; Score 301; DB 4; Length 524;
Best Local Similarity 25.2%; Pred. No. 2.5e-22;
Matches 121; Conservative 66; Mismatches 151; Indels 142; Gaps 20;
QY 115 WLQTLREMLEK-----LTSRPHKLLVPINPFQGGKQKRIYERKVAPLFTLASIT 164
Db 17 WYRSURWQLHRTLEIFVAPTVDERRRRVLLNPNKSGSGDAREVFNNMHVTPVLNEARVP 76
QY 165 TDIIVTEHANOAKETLYEINIDKYGVCGDGMFSEVLHGLIGRTQRSAGVDQNHPR 224
Db 77 YDLYTKNSFAIFELSTRCLDAMCCVAVGDLGFHEIVNGLLRQ-----DWAH--- 127
QY 225 VLVPSSLRIGIIPAGSTDCV-----CYS--TVGTSDAETSALHIVVGDSLAMDVSSVHH 276
Db 128 --VLPHLALGIIPCGSGNGLARSIAHCYNKPVLG-----AALTIVSGRSSPMVVRVOL 179
QY 277 NSTLRYSVLLGYGYFDIHKDSEKKRWGLARYDFGLKFTFLSHHCYGVTVSFLPAQH 336
Db 180 QSRSL-YSFSLSGWGLISDVDESIRMLGYQRFVTMTVLRLNLTNGRISYLLTDH 238
QY 337 TVGSPRD-----RKPCRACF-----VCROSKOOLEE----- 364
Db 239 EVSSTHSATGYAAQRMRQSSRCNTHIDMLNGPAPIYHSSAEYLPQEFADVISLETSLNQ 298
QY 365 -----OKKALY-----GLEAAE-----DV 378
Db 299 SFRSRCDMSLGGSRRSFYYSISESIYHSLADESEFAGLAASLENROQNYGPASELPDL 358
QY 379 EE-----WQVVCCKFLAINA---TNMSCACRRSPRGLSPAHLGDGSSDLILIRK-CS 427
```

```
Db 359 NEPLUSEDOGWLVEGEFVMMHAVYQTHLGDCH-----FAPKAQNDGTIVLILIRAGIS 413
QY 428 RFNFLRFLIR---HTNQODQDFTFEVEYRVYKFKQFTSKHMEDESD---LKEGKKRFG 481
Db 414 RPHLLSFLYNSSGTHLPESH-DHVKVLPVRAFR-----LEPYDNHGIIITVDGERVEFG 467

RESULT 6
US-09-970-516-4
; Sequence 4, Application US/09970516
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-516-4

Query Match 9.4%; Score 271.5; DB 4; Length 618;
Best Local Similarity 30.0%; Pred. No. 4e-19;
Matches 78; Conservative 44; Mismatches 111; Indels 27; Gaps 4;
QY 91 RARHRWKAAQVTFWCPEEQQLCHLWLOTLRMLBKLTSRPHKLLVFINPFGGKGGKRIY 150
Db 111 RAEQRWATALTCL-----LRGLPLGDSGITPDLPRPPRLLLLVNPFQGRGLAWQWC 164
QY 151 ERKVAPLFTLASITTDIIVTEHANOAKETLYEINIDKYGVCGDGMFSEVLHGLIGR 210
Db 165 KNHVLPMISEAGLSFNLIQTERQHARELVQGLSEMDGIVTVSGDGLLHEVLNGLLDR 224
QY 211 TORSAGVDQNHPRAVLPSSLRIGIIPAGSTDCVCYS-----TVGTSDAETSALHI 261
Db 225 -----PDWEAEVKNP-----VGLPCGSGNALAGAVNQHGFEPALEGDLNLCSSLL 272
QY 262 VGDSLAMDVSSVHHNSTLLRYSVLLGYGYFDIHKDSEKKRWGLARYDFSGLKTFLS 321
Db 273 CRGGHPDLILLSVTLASGRCFSLSVAWGFVSDVIQSERFRAIGSARFTLTGLVGLAT 332
QY 322 HHCYEGTVSFLPAQHTVGSP 341
Db 333 LHTRYGRLSYLPATVPEPASP 352

RESULT 7
US-09-817-676A-14
; Sequence 14, Application US/09817676A
; Patent No. 6800470
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817,676A
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/194,318
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin ver. 2.0
; SEQ ID NO 14
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-817-676A-14
```

Query Match	9.4%	Score	271.5	DB	4	Length	618
Best Local Similarity	30.0%	Pred. No.	4e-19				
Matches	78	Conservative	44	Mismatches	111	Indels	27
Gaps							
Qy	91	RARRHRKWAQVTPWCPEEQI	CHLWQLT	LRMLEKLT	SRPKHLV	FINPFGGKGQ	KRIY 150
Db	111	RAEAQRWATALTCL	-----LRGLPL	PCDGEIT	PDLLPR	PRLLLVN	PFGGRGLAWQC 164
Qy	151	ERKVAPLFTLASITTD	IIVTEHANQAKET	LYEINID	KVDGIV	CVCGDMGF	SEVLHGLIGR 210
Db	165	KNHVLPMISEAGLSFNLI	QIOTERQHARHEL	VGGLSE	WDGIVT	VGSGGLL	HEVLNGLLDR 224
Qy	211	TQRSAGVDQNHPR	AVLPSSLRIGII	I	PAGSTDC	VCYS-----	TVGTSDAETSALHI 261
Db	225	-----PDWEAEAVKMP	-----	VGILPC	SGSNALAGA	VNQHG	GFEPALGLDLLNCSLL 272
Qy	262	VVGDSLAMDYSSV	HHNSTLLRYS	VSLLG	YGFVDII	KDSEKKW	IGLARYDFSGIKTFLS 321
Db	273	CRGGGHPLDLLSV	TLASGSR	CFSLF	VAWGF	SVDDIQ	SERFALGARSFTLGTIVGLAT 332
Qy	322	HHCYEGTVSFL	PAQHTV	GSP			341
Db	333	LHTYGRGLSYL	PATVE	PASP			352

RESULT 8  
 US-09-817-676A-12  
   ; Sequence 12, Application US/09817676A  
   ; Patent No. 6800470  
   ; GENERAL INFORMATION:  
   ; APPLICANT: Spiegel, Sarah  
   ; APPLICANT: Kohama, Takafumi  
   ; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,  
   ; TITLE OF INVENTION: Expression and Methods of Use Thereof  
   ; FILE REFERENCE: 00170/HG  
   ; CURRENT APPLICATION NUMBER: US/09/817,676A  
   ; CURRENT FILING DATE: 2001-03-26  
   ; PRIOR APPLICATION NUMBER: US 60/194,318  
   ; PRIOR FILING DATE: 2000-04-03  
   ; NUMBER OF SEQ ID NOS: 15  
   ; SOFTWARE: PatentIn Ver. 2.0  
   ; SEQ ID NO 12  
   ; LENGTH: 617  
   ; TYPE: PRT  
   ; ORGANISM: Mus musculus  
 US-09-817-676A-12

Query Match	9.3%	Score	267.5;	DB	4;	Length	617;
Best Local Similarity	28.5%;	Pred.	No. 1e-18;				
Matches	79;	Conservative	49;	Mismatches	122;	Indels	27; Gaps
Qy	67	GKHOGSGKQWKEPYAFTHCVRARRHRWKVAQTFCPEEQCLCHLWLQTLREMLEKL	126	:::	:::	:	: ::
Db	88	GRGRRRATRTFRADGATTVEENRAEAQRWATALTCL-----LRGVPLSGDQEITPEL	141	:	::	:	::
Qy	127	TSRPKHLVFNPPGGKGQKRIVYRKVAPLFTLASITTDIIIVTEHANOAKETIYEIND	186	:::		:	: ::
Db	142	LPRKPRLLILNPPGGRGLAWQRCDMHDVPMISAGLSFNLIQTERQHABELVOGLSL	201	:::		:	: ::
Qy	187	KYDGIVCVGGDMPSFVLHGILIGRTORSAGVDQNHPRAVLPVSSLRIGIIPAGSTDCVC-	245	:::		:	: ::
Db	202	EWEGIVTVSGDLLYEVLNGLLDR-----PDWEDAVRMP----IGVLPFGSGNALAG	249	:::		:	: ::
Qy	246	-----YSTVCTSDAETSALHIWVGDSLAMDVSSVHHNSTILARYSVSLGCGFYGDII	297	:::		:	: ::
Db	250	AVSHHGFEQVGVDDLLINCSLLRCRGSGHPDLDSLVTLASGSRCSFSLVAMWGFLSDVD	309	:::		:	: ::
Qy	298	KDSKKXWLGRLARDYDFGLKTFLSHHCYEGTVSFPLA	334	:::		:	: ::
Db	310	IHSRRFRALGSARFTTGAVLGLASLHTYGRSLYLPA	346	:::		:	: ::

```

RESULT 9
US-09-959-897-2
; Sequence 2, Application US/09959897
; Patent No. 6730480
;
; GENERAL INFORMATION:
; APPLICANT: PITSON, Stuart M
; APPLICANT: Brian, WATTENBERG W
; APPLICANT: Pu, XIA
; APPLICANT: Richard, D'ANDREA J
; APPLICANT: Jennifer, BAMBLE R
; APPLICANT: Mathew, VADAS A
;
; TITLE OF INVENTION: SPHINGOSINE KINASE ENZYME
; FILE REFERENCE: PITSON=1
; CURRENT APPLICATION NUMBER: US/09/959,897
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: PCT/AU00/00457
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: AU PQ 0339
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: AU PQ 1504
; PRIOR FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-959-897-2

```

Query Match	9.2%;	Score 266.5;	DB 4;	Length 384;
Best Local Similarity	25.3%;	Pred. No. 6.3e-19;		
Matches 107;	Conservative	68;	Mismatches 169;	Indels 79; Gaps 16;
Qy	129	RPKHLVFINPFGGKGKRIYERKVAPLFTLASITDDIIVTEHANAKETLYEINIDKY	188	
Db	13	RPCEVLVLLNPRCGKGKALQLFRSHVQPLLAEEAIEFTLMITERNHARELVRSEELGRW	72	
Qy	189	DGIVCVGGDGMFSEVLHGLIGRTORSAGVQDNHPRAVLVPSSLRIGHIIPAGSTDCVC---	245	
Db	73	DALVMSGDGLMEHVNGLMERPDWETAIOK-----PLCSLPAGSGNALAASL	120	
Qy	246	----YSTVGTSGDAETSAHSHVVGDSL-A-MDVSSVHHNSTLLRYSVSLGLGYFGFYDIKD	299	
Db	121	NHVAGYEQVTNEDLLTNTCTILLCERLLSPMNLLSLHTASGLRFLSVLSLAWGFTADVDLE	180	
Qy	300	SEKRWGLGLARYDPFSGUKTFLSHHCYEGTVSFPLPAQHTVGSPPDRKPCRCAGCFVCRQSKQ	359	
Db	181	SEKRYRLGEKRFITLGTFLRLAALTFRGLAYLFPVGR-VGSKTPASP-----VVVQOGPV	234	
Qy	360	QLEBEQKALYGLAEAAEDV-EEMOVVCGK---FLAINATNMSCACRRSPRLGSPAALHG	414	
Db	235	DAH-----LVPLE--RPVPSHTVVDEDFVLVALLHSHLGSEMFAAPMGRCAA-----	282	
Qy	415	DGSSDLILIRK-CSRFNFLR-FLIRHTNQOQDFDTFVYVRVKFOFTSKHMEDESDSL	472	
Db	283	-GVMHFLFYRAGVSRAMLRLFLAMEKGRHMEYECPLYLVVPVVAFR-----LEPKD---	333	
Qy	473	KEGKKKFGHICSSHPSCCCTVSNSWNCDGEVLHSPAIEVRVHCQLVRLFARGIEENP-	531	
Db	334	---GRGMFA-----VDGELMVSEAVQGVHPNTVFMVWGCVGCVPEPPS	371	
Qy	532	-KP	533	
Db	372	WKP	374	

RESULT 10  
US-09-970-516-2  
; Sequence 2, Application US/09970516  
; Patent No. 6610534  
; GENERAL INFORMATION:  
; APPLICANT: No. 6610534artis AG  
; TITLE OF INVENTION: Induction of b

; TITLE OF INVENTION: polynucleotides encoding sphingosine kinases

```
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-516-2

Query Match          9.2%; Score 264.5; DB 4; Length 384;
Best Local Similarity 25.3%; Pred. No. 1e-18;
Matches 107; Conservative 68; Mismatches 169; Indels 79; Gaps 16;

QY 129 RPKHLVFINPFGGKQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETLYEINIDKY 188
DB 13 RPCRVLVLLNPRGGKQKALQFRSHVQPLLAEEISFTLMLTERNHARELVRSEELGRW 72

QY 189 DGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGSTDCVC--- 245
DB 73 DALVMSGDGLMHEVVGMLMERPDWETAIQK-----PLCSLPAGSGNALAASL 120

QY 246 -----YSTVGTSDAETSALHIVVGDLSA-MDVSSVHHNSTLLRYSVSLGFGYCDI1KD 299
DB 121 NHYAGYEQVTNEDLLTNTCLLCRRLLSPMNLISLTASGLRFSVLGLAWGFIADVDLE 180

QY 300 SEKKRWGLARYDFSLKTFLSHHCYEGTVSFLPAQHTVGSPRDKPCRAGCFVCQSQK 359
DB 181 SEKYRRLGEMRTLTGTFRLAALRYRGLAYLPVGR-VGSKTPASP-----VVVQGPV 234

QY 360 QLEEEQKALYGLEAAEDV-EEWQVVCVK-----FLAINATNMSCACRRSPRGLSPAHLG 414
DB 235 DAH-----LVPLE--EPVSHWTVPDEDFVLVLLALLHSLGSEMFAPMGRCAA--- 282

QY 415 DGSSDLILIRK-CSRFNPLRLI-----RHTNQDQDFTFVEVYRVKKFQFTSKHMEDE 472
DB 283 -GVMEHLFYVRAGVSRAMLRLFLAMEKGRHM-----EYECPLYVYVPVAVFR-----LEPKD--- 333

QY 473 KEGGKRRFGHICSSHPSCCTVSNSSWNCDEVLHSPAIEVVRVHCOLVRLPARGIENP- 531
DB 334 ---GKGVFA-----VDGELMVSEAVQGVQVHPNFMVMSGCVPEPPS 371

QY 532 -KP 533
DB 372 WKP 374
```

```
RESULT 11
US-09-949-016-7026
; Sequence 7026, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7026
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7026
```

```
Query Match          9.1%; Score 263.5; DB 4; Length 384;
Best Local Similarity 25.1%; Pred. No. 1.3e-18;
Matches 107; Conservative 69; Mismatches 164; Indels 87; Gaps 17;

QY 129 RPKHLVFINPFGGKQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETLYEINIDKY 188
DB 13 RPCRVLVLLNPRGGKQKALQFRSHVQPLLAEEISFTLMLTERNHARELVRSEELGRW 72

QY 189 DGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGSTDCVC--- 245
DB 73 DALVMSGDGLMHEVVGMLMERPDWETAIQK-----PLCSLPAGSGNALAASL 120

QY 246 -----YSTVGTSDAETSALHIVVGDLSA-MDVSSVHHNSTLLRYSVSLGFGYCDI1KD 299
DB 121 NHYAGYEQVTNEDLLTNTCLLCRRLLSPMNLISLTASGLRFSVLGLAWGFIADVDLE 180

QY 300 SEKKRWGLARYDFSLKTFLSHHCYEGTVSFLPAQHTVGSPRDKPCRAGCFVCQSQK 359
DB 181 SEKYRRLGEMRTLTGTFRLAALRYRGLAYLPVGR-VGSKTPASP-----VVVQGPV 234

QY 360 QLEEEQKALYGLEAAEDV-EEWQVVCVK-----FLAINATNMSCACRRSPRGLSPAHLG 414
DB 235 DAH-----LVPLE--EPVSHWTVPDEDFVLVLLALLHSLGSEMFAPMGRCAA--- 282

QY 415 DGSSDLILIRK-CSRFNPLRLI-----RHTNQDQDFTFVEVYRVKKFQFTSKHMEDE 468
DB 283 -GVMEHLFYVRAGVSRAMLRLFLAMEKGRHM-----EYECPLYVYVPVAVFR-----LEPK 332

QY 469 DSDLKEGKRGHICSSHPSCCTVSNSSWNCDEVLHSPAIEVVRVHCOLVRLPARGIE 528
DB 333 D-----GKGVFA-----VDGELMVSEAVQGVQVHPNFMVMSGCV 367

QY 529 ENP--KP 533
DB 368 PPSWKP 374

RESULT 12
US-09-796-487-3
; Sequence 3, Application US/09796487
; Patent No. 6830916
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
; FILE REFERENCE: 07320001aa (2033957-0001)
; CURRENT APPLICATION NUMBER: US/09/796,487
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,532
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: US 09/530,868
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(384)
; OTHER INFORMATION: seq id no 3 is the peptide sequence of hSPHK1 in Fig. 3, correspo
; nding to amino acid residue 1 to 384 of Homo sapiens SPHK-1 of Ge
; nBank sequence Accession Number AAF73423.
; PUBLICATION INFORMATION:
; AUTHORS: Nava et al.
; TITLE: Functional characterization of human sphingosine kinase-1
; JOURNAL: FEBS Lett.
; VOLUME: 473
; ISSUE: 1
; PAGES: 81-84
; DATE: 2000
; DATABASE ACCESSION NUMBER: AAF73423
```



```

; ISSUES: 37
; PAGES: 23722-23728
; DATE: 1998
; DATABASE ACCESSION NUMBER: AAC61698
; DATABASE ENTRY DATE: 1998-09-26
; RELEVANT RESIDUES: (1)..(388)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AAC61698
; DATABASE ENTRY DATE: 1998-09-26
; RELEVANT RESIDUES: (1)..(388)
US-09-796-487-2

Query Match
Best Local Similarity 23.3%; Pred. No. 3.1e-18; Length 388;
Matches 100; Conservative 70; Mismatches 173; Indels 86; Gaps 13;

Qy 104 FWCPEQLCHLWQLTREMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTLASI 163
Db 2 WWC-----CVLFVV---ECPRGLLPFCRVLVLLNPQGGKALQLFQSRVQPFLEAEI 53

Qy 164 TTDIIIVTEHANOAKETLYEINIDKYDGIYVCGDGMFSEVLHGLIGRTQRSAGVDQNHPR 223
Db 54 TFKLILTKRKHARELVCAEELGHWDALAVMSGDGLMHEVVNGLMERPDWETAIOK---- 109

Qy 224 AVLVPSSLRIGIIPAGSTDCVC-----YSTVGTSDAETSALHIVVGDLSA-MDYSSV 274
Db 110 -----PLCSLPGSGNALAASVNHVAGYEQVTNEDLLINCTLLCCRRLLSPMNLSSL 161

Qy 275 HNNSTLLRYSVLLGYGYGDIIDSEKKRWGLARYDFGLKTFSLHHCYEGTVSFLPA 334
Db 162 HTASGLRLYSVLSWGFVADVLESEKYRRLGEIRFTVGTFFRLASRLRYQGLAYLP- 220

Qy 335 QHTVGSPRDRKPCRAGCFVCRSQKQLEEBEOKKALYGLEAAEDV-BEQVVVCGK----FL 389
Db 221 ---VGTVASKRPA-----STLVQKGPVDTHLVPLEEPVPSHTVTVPEQDFVLVL 266

Qy 390 AINATNSCACRRSPRGLSPAHLGDSGLDSSDLILIRK-CSRFNFLR-FLIRHTNQDQDFP 447
Db 267 VLLHTLHLSSELFAAPMGRCFA-----GYMHLFYVRAGVSRALLRLFLAMQKGMELDC 321

Qy 448 TFVEYRVKQFTSKHMEDESDLEKGGKKRFGHICSSHPSCCCTVSNSSWNCDEVLH 507
Db 322 PYLVHVPVVAFLRLEPRS-----ORGFVSDGELMV 351

Qy 508 SPAIEVRVH 516
Db 352 CEAVQGVH 360

```

## RESULT 15

```

US-10-053-510-21
; Sequence 21, Application US/10053510
; Patent No. 6830881
; GENERAL INFORMATION:
; APPLICANT: Saba, Julie D.
; APPLICANT: Fyrest, Henrik
; TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
; TITLE OF INVENTION: POLYNUCLEOTIDES AND MODULATING AGENTS AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 200116.402C2
; CURRENT APPLICATION NUMBER: US/10/053.510
; CURRENT FILING DATE: 2002-01-17
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-053-510-21

```

```

Query Match
Best Local Similarity 25.1%; Pred. No. 8.4e-18; Length 368;
Matches 105; Conservative 68; Mismatches 167; Indels 79; Gaps 16;

```

```

Qy 133 LLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLYEINIDKYDGI 192
Db 1 VLVLLNPRGGKALQLFRSHVQPLLAETISFTMLTERRNHARELVRSEELGRWDALV 60

Qy 193 CVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGSTDCVC----- 245
Db 61 VMSGDGLMHEVVNGLMERPDWETAIOK-----PLCSLPGSGNALAASLHNYA 108

Qy 246 -YSTVGTSDAETSALHIVVGDLSA-MDYSSVHHNSTLLRYSVLSLIGYGYGDIIDSEKK 303
Db 109 GYEQVTNEDLLTNCTLLCCRRLLSPMNLSSLHTASGLRLFSVLSLANGFIADVLESEKY 168

Qy 304 RWLGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCRSQKQLE 363
Db 169 RRLGEMRFTLTGTLRLAALRTYRGLAYLPVGR-VGSKTPASP-----VVVQGPVDAH- 221

Qy 364 EQKKALYGLEAAEDV-BEQVVVCGK-----FLAINATNSCACRRSPRGLSPAHLGDS 418
Db 222 -----LVPLE--EPVPSHTVVPDEDFVLVLLHLSHLSGEMFAAPMGRCFA-----GVM 269

Qy 419 DLILIRK-CSRFNFLR-FLIRHTNQDQDFTFVEYRVKQFTSKHMEDESDLEKGG 476
Db 270 HLFYVRAGVSRALLRLFLAMEKGRHMEYECPLYLVYVVAFR-----LEPKD-----G 318

Qy 477 KKRFGHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENP--KP 533
Db 319 KGVEA-----VDGELMVSEAVQGVHPNYFMWVSGCVPEPPPSWKP 358

```

Search completed: September 3, 2005, 04:10:36  
Job time : 30.3937 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (C) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model  
Run on: September 3, 2005, 04:09:18 ; Search time 109.661 Seconds  
(without alignments)  
1893.930 Million cell updates/sec

Title: US-10-631-958-10  
Perfect score: 2888  
Sequence: 1 NGATGAAPLQSLVWVKKQR.....QLVRLFARGIENKPDPSHS 537

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A: Geneseq 16Dec04: \*  
1: Geneseqp1980s: \*  
2: Geneseqp1990s: \*  
3: Geneseqp2000s: \*  
4: Geneseqp2001s: \*  
5: Geneseqp2002s: \*  
6: Geneseqp2003as: \*  
7: Geneseqp2003bs: \*  
8: Geneseqp2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2888	100.0	537	ABB07856	Abb07856 Human sph
2	2888	100.0	537	ADA05880	Ada05880 Human NOV
3	2888	100.0	537	ADJ96664	Adj96664 Human lip
4	2888	100.0	537	ADN62845	Adn62845 Human NOV
5	2888	100.0	562	ABB07857	Abb07857 Human sph
6	2880	99.7	537	AAM49115	Aam49115 Human Cer
7	2588	89.6	481	ABR56302	AbR56302 Human Sph
8	2463	85.3	460	AAY96059	Aay96059 Human sph
9	2456.5	85.1	471	AAE07884	Aae07884 Human sph
10	2428	84.1	531	ADP55248	Adp55248 Human PRO
11	2350	81.4	536	ADS11054	Ads11054 Human the
12	2210	76.5	416	ABR56301	AbR56301 Human Sph
13	2208.5	76.5	746	ABG13541	Abg13541 Novel hum
14	1714.5	59.4	727	ABG13543	Abg13543 Novel hum
15	1640.5	56.8	326	ABB07854	Abb07854 Human sph
16	1361.5	47.1	454	ABG13544	Abg13544 Novel hum
17	1055	36.5	228	ABAB1822	Aab1822 Human ORF
18	1035	35.8	255	ADS12267	Ads12267 Human the
19	1032	35.7	190	ABP64913	Abp64913 Human pro
20	986	34.1	182	ABA42383	Aba42383 Human ORF
21	675	23.4	596	ABB69669	Abb69669 Drosophil
22	664	23.0	136	ABG13540	Abg13540 Novel hum
23	647	22.4	144	AAE07885	Aae07885 Partial r
24	628	21.7	136	ABG13542	Abg13542 Novel hum
25	582.5	20.2	532	ADQ88891	Adq88891 Novel hum

26	335	11.6	79	4	AAE07886	Aae07886 Partial m
27	310.5	10.8	641	4	ABB58465	Abb58465 Drosophil
28	310.5	10.8	641	6	ABR82392	AbR82392 D. melano
29	306	10.6	490	6	ABR82389	AbR82389 D. melano
30	301	10.4	524	6	ABR82390	AbR82390 D. melano
31	301	10.4	907	4	ABB57980	Abb57980 Drosophil
32	301	10.4	907	6	ABR82393	AbR82393 D. melano
33	277.5	9.6	618	4	AAU09075	Aau09075 Human sph
34	276	9.6	806	4	ABG21144	Abg21144 Novel hum
35	271.5	9.4	618	4	AAU09074	Aau09074 Human sph
36	271.5	9.4	618	5	ABG31587	Abg31587 Human sph
37	271.5	9.4	618	5	ABB07855	Abb07855 Protein s
38	271.5	9.4	618	7	ADE38417	AdE38417 Human can
39	271.5	9.4	618	8	ADQ15182	Adq15182 Human can
40	271.5	9.4	654	4	ABU52806	Abu52806 Human sig
41	271.5	9.4	654	8	ADJ66569	Adj66569 Sphingos
42	267.5	9.3	617	4	AAU09073	Aau09073 Mouse sph
43	266.5	9.2	384	4	AAB48007	Aab48007 Human sph
44	266.5	9.2	384	7	ADF28783	Adf28783 Human sph
45	264.5	9.2	384	3	AAB18659	Aab18659 A human r

ALIGNMENTS

RESULT 1

ABB07856	ABB07856 standard; protein; 537 AA.
XX	AC ABB07856;
XX	AC ABB07856;
DT	03-JUL-2002 (first entry)
XX	XX
DE	Human sphingosine kinase-like protein.
XX	XX
KW	Human sphingosine kinase-like protein; intracellular signalling;
KW	cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;
KW	autoimmune disease; rheumatoid arthritis; Parkinson's disease.
XX	XX
OS	Homo sapiens.
XX	XX
PN	WO200228906-A2.
XX	XX
PD	11-APR-2002.
XX	XX
PF	05-OCT-2001; 2001WO-EP011516.
XX	XX
PR	06-OCT-2000; 2000US-0238005P.
XX	23-AUG-2001; 2001US-0314113P.
XX	(FARB ) BAYER AG.
XX	Kossida S, Encinas J;
XX	WPI; 2002-340094/37.
DR	N-PSDB; ABL40828.
DR	XX
PT	New reagent for modulating the activity of sphingosine kinase-like
PT	protein polypeptide or polynucleotide and treating cancer, asthma,
PT	allergy, an autoimmune disease, or a central or peripheral nervous system
PT	disorder.
XX	XX
PS	Claim 25; Fig 10; 120pp; English.
XX	XX
CC	The invention relates to a human sphingosine kinase-like protein. The
CC	polypeptide can be expressed by standard recombinant methodology. The
CC	sphingosine kinase-like protein and gene can be used to regulate
CC	intracellular signalling and consequently cell proliferation and
CC	apoptosis. Such regulation is useful for treating cancer, allergies (e.g.
CC	asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and
CC	peripheral nervous system disorders (e.g. Parkinson's disease). The
CC	present sequence represents the human sphingosine kinase-like protein
XX	XX

sq	Sequence 537 AA;	
Query Match	100.0%; Score 2888; DB 5; Length 537;	
Best Local Similarity	100.0%; Pred. No. 2e-274;	
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1 MGATGAEPLOSVLVWKQORCAVSLERPARALLRWRSFPGAGAPGADACSVFVSEIIAV 60	
Db	1 MGATGAEPLOSVLVWKQORCAVSLERPARALLRWRSFPGAGAPGADACSVFVSEIIAV 60	
Qy	61 EETDVHGHQSGKQKMEKEFYAFTVHCVKARHRWKWAQVTFWCPEEQQLCHLWLQTLR 120	
Db	61 EETDVHGHQSGKQKMEKEFYAFTVHCVKARHRWKWAQVTFWCPEEQQLCHLWLQTLR 120	
Qy	121 EMLEKLTSRPKHLVFINPFGKGQGRKIYERKVAPLFTLASITTDIIIVTEHANOAKETL 180	
Db	121 EMLEKLTSRPKHLVFINPFGKGQGRKIYERKVAPLFTLASITTDIIIVTEHANOAKETL 180	
Qy	181 YEINIDKDVICVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGIIPAGS 240	
Db	181 YEINIDKDVICVGGDMFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGIIPAGS 240	
Qy	241 TDCVCYSTVGTSDAETSALHIVVGDSLAMDVSSVHHNSTLLRYSVSLLGFGFYGDIIDKS 300	
Db	241 TDCVCYSTVGTSDAETSALHIVVGDSLAMDVSSVHHNSTLLRYSVSLLGFGFYGDIIDKS 300	
Qy	301 EKKRWGLIARVDFSLKTFSLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGFCVCRSQKQ 360	
Db	301 EKKRWGLIARVDFSLKTFSLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGFCVCRSQKQ 360	
Qy	361 LEEQKALYGLEAAEDVEEVQVCGKFLAATNATMNSCACRRSPRGLSPAHLGDGSSDL 420	
Db	361 LEEQKALYGLEAAEDVEEVQVCGKFLAATNATMNSCACRRSPRGLSPAHLGDGSSDL 420	
Qy	421 ILIRKCSFNLFLRIHTNQDQDFTFVEVYRVKKFQFTSKMEDESDSLKGGKKRF 480	
Db	421 ILIRKCSFNLFLRIHTNQDQDFTFVEVYRVKKFQFTSKMEDESDSLKGGKKRF 480	
Qy	481 CHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVFLFARGTEENPKPDSHS 537	
Db	481 CHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCOLVFLFARGTEENPKPDSHS 537	

RESULT 2	
ADA05680	
ID	ADA05680 standard; protein; 537 AA.
XX	
AC	ADA05680;
XX	
DT	06-NOV-2003 (first entry)
XX	
DE	Human NOV9a protein SEQ ID NO:40.
XX	
KW	human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
KW	immunomodulator; cytosolic; nootropic; neuroprotective;
KW	antiparkinsonian; antilipaemic; gene therapy; human disease;
KW	metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
KW	neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW	immune disorder; haematopoietic disorder; dyslipidaemia.
XX	
OS	Homo sapiens.
XX	
FN	WO200029424-A2.
XX	
PD	10-APR-2003.
XX	
PF	02-OCT-2002; 2002WO-US031373.
XX	
PR	02-OCT-2001; 2001US-0326483P.
PR	05-OCT-2001; 2001US-0327435P.
PR	05-OCT-2001; 2001US-0327449P.
PR	09-OCT-2001; 2001US-0327917P.
PR	09-OCT-2001; 2001US-0328029P.
PR	09-OCT-2001; 2001US-0328044P.
PR	09-OCT-2001; 2001US-0328056P.
PR	12-OCT-2001; 2001US-0328849P.
PR	15-OCT-2001; 2001US-0329414P.
PR	17-OCT-2001; 2001US-0330142P.
PR	18-OCT-2001; 2001US-0330309P.
PR	22-OCT-2001; 2001US-0341058P.
PR	24-OCT-2001; 2001US-0339266P.
PR	24-OCT-2001; 2001US-0343629P.
PR	29-OCT-2001; 2001US-0349575P.
PR	01-NOV-2001; 2001US-0346357P.
PR	17-APR-2002; 2002US-0373260P.
PR	19-APR-2002; 2002US-0373815P.
PR	19-APR-2002; 2002US-0373817P.
PR	19-APR-2002; 2002US-0373826P.
PR	19-APR-2002; 2002US-0373884P.
PR	22-APR-2002; 2002US-0374977P.
PR	16-MAY-2002; 2002US-0381037P.
PR	16-MAY-2002; 2002US-0381038P.
PR	17-MAY-2002; 2002US-0381042P.
PR	17-MAY-2002; 2002US-0381642P.
PR	28-MAY-2002; 2002US-0383656P.
PR	29-MAY-2002; 2002US-0383831P.
PR	25-JUN-2002; 2002US-0391335P.
PR	01-OCT-2002; 2002US-00262511.
XX	
XX	(CURA-) CURAGEN CORP.
PA	Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
PI	Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
PI	Ort T, Gorman L, Zexhusen BD, Anderson DW, Catterton E;
PI	Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;
PI	Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, DiPippo VA;
PI	Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;
XX	
XX	WPI; 2003-381626/36.
DR	N-PSDB; ADA05679.
XX	
PT	New NOVX polypeptides and nucleic acids, useful for diagnosing,
PT	preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
PT	cancer or dyslipidemia, and in chromosome mapping, tissue typing or
PT	pharmacogenomics.
XX	
PS	Claim 1; Page 135; 586pp; English.
XX	
CC	The present invention describes NOVX proteins, where X can be 1 to 55
CC	(e.g. NOV1). Also described: (1) a composition comprising a polypeptide
CC	described above and a carrier; (2) a kit comprising, in one or more
CC	containers, the composition described above; (3) an isolated nucleic acid
CC	molecule which encodes a NOVX protein of the invention; (4) a vector
CC	comprising the nucleic acid molecule described above; (5) a cell
CC	comprising the above vector; (6) an antibody that immunospecifically
CC	binds to the polypeptide described above; (7) methods for determining the
CC	presence or amount of the above polypeptide or nucleic acid molecule in a
CC	sample; (8) methods for determining the presence of or predisposition to
CC	a disease associated with altered levels of expression of the above
CC	polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
CC	method of identifying an agent that binds to the polypeptide described
CC	above; (10) a method for identifying a potential therapeutic agent for
CC	use in treating a pathology that is related to an aberrant expression or
CC	aberrant physiological interactions of the polypeptide; (11) a method of
CC	screening for a modulator of activity or of latency or predisposition to
CC	a pathology associated with the polypeptide; (12) a method for modulating
CC	the activity of the polypeptide described above; (13) methods of treating
CC	or preventing a pathology associated with the above polypeptide in a
CC	mammal; and (14) a method for producing the above polypeptide. NOVX
CC	sequences have antidiabetic, anorectic, antibacterial, virucide,
CC	immunomodulator, cytosolic, nootropic, neuroprotective, antiparkinsonian
CC	and antilipaemic activities, and can be used in gene therapy. The
CC	polypeptide is useful in manufacturing a medicament for treating a
CC	syndrome associated with a human disease. The polypeptide or the nucleic
CC	acid molecule may be used to diagnose, treat or prevent metabolic
CC	disorders such as diabetes or obesity, infections, cachexia, cancer,

```
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
CC disease, immune disorders, haematopoietic disorders and various
CC dyslipidaemias. The nucleic acids can also be used as hybridisation
CC probes, in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. The present sequence represents a human NOVX from the
CC present invention.
XX
SQ Sequence 537 AA;

Query Match      100.0%; Score 2888; DB 6; Length 537;
Best Local Similarity 100.0%; Pred. No. 2e-274;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGATGAAPLOSIVLVKQRCVSLPARALLRWRSPGAGAGADACSVPSVEITAV 60
Db 1 MGATGAAPLOSIVLVKQRCVSLPARALLRWRSPGAGAGADACSVPSVEITAV 60

Qy 61 BETDVHGKHGSGKWKQKQKPYAFTVHCVRARRRHWKWAQVTFWCPEQLCHLWLOTLR 120
Db 61 BETDVHGKHGSGKWKQKQKPYAFTVHCVRARRRHWKWAQVTFWCPEQLCHLWLOTLR 120

Qy 121 EMLEKLTSRPKHLLVFINPFGKGQGGKRIYERKVAPLFTLASITTDIIIVTEHANOAKETL 180
Db 121 EMLEKLTSRPKHLLVFINPFGKGQGGKRIYERKVAPLFTLASITTDIIIVTEHANOAKETL 180

Qy 181 YEINIDKYDGIIVCGDGMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPAGS 240
Db 181 YEINIDKYDGIIVCGDGMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPAGS 240

Qy 241 TDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGAGYFGDIIKDS 300
Db 241 TDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGAGYFGDIIKDS 300

Qy 301 EKRWGLGARYDFSGDKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGFCVCRQSKQ 360
Db 301 EKRWGLGARYDFSGDKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGFCVCRQSKQ 360

Qy 361 LEEBQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 420
Db 361 LEEBQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 420

Qy 421 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRKQFTSKHMEDESDSKEGGKRF 480
Db 421 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRKQFTSKHMEDESDSKEGGKRF 480

Qy 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 537
Db 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 537

RESULT 3
ID ADJ96664
XX ADJ96664 standard; protein; 537 AA.
XX
AC ADJ96664;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human lipid kinase KIAA1646 protein SeqID 121.
XX
KW kinase; human; tyrosine protein kinase; serine/threonine protein kinase;
KW PTK; STK; gene therapy; cancer; immune-related disease;
KW cardiovascular disease; brain; neuronal associated disease; metabolic;
KW inflammatory disorder; cytostatic; neuroprotective; immunomodulator;
KW antiinflammatory; enzyme; lipid kinase; KIAA1646.
XX
OS Homo sapiens.
XX 44.
XX WO2004006838-A2.
XX
PD 22-JAN-2004.
XX
```

```
PF 15-JUL-2003; 2003WO-US021730.
XX
PR 15-JUL-2002; 2002US-0395632P.
XX
PA (SUGE-) SUGEN INC.
XX
PI Whyte D, Manning G, Caenepeel S;
XX
DR WPI; 2004-122753/12.
XX
N-PSDB; ADJ96598.
XX
PT New nucleic acid molecule encoding a kinase polypeptide, useful for
PT preparing a composition for treating diseases or disorders, e.g., cancer,
PT or neurological, immunological or inflammatory disorders.
XX
PS Claim 1; SEQ ID NO 121; 366pp; English.
XX
CC This invention relates to a novel isolated, enriched or purified nucleic
CC acid molecule that encodes a kinase polypeptide. Specifically, it relates
CC to human tyrosine and serine/threonine protein kinases (PTK's and STK's),
CC as well as protein kinase-like enzymes. The present invention describes
CC screening methods to identify agonists, antagonists and antibodies that
CC can be used to modulate the activity or function of the mammalian kinase
CC enzymes. As such, these compositions can be used for gene therapy
CC purposes to treat diseases or disorders including cancer, immune-related
CC diseases, cardiovascular disease, brain or neuronal associated disease,
CC metabolic and inflammatory disorders. Accordingly, they exhibit
CC cytosatic, neuroprotective, immunomodulator and antiinflammatory
CC activities. This polypeptide sequence is a human kinase protein sequence
CC of the invention.
XX
SQ Sequence 537 AA;

Query Match      100.0%; Score 2888; DB 8; Length 537;
Best Local Similarity 100.0%; Pred. No. 2e-274;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGATGAAPLOSIVLVKQRCVSLPARALLRWRSPGAGAGADACSVPSVEITAV 60
Db 1 MGATGAAPLOSIVLVKQRCVSLPARALLRWRSPGAGAGADACSVPSVEITAV 60

Qy 61 BETDVHGKHGSGKWKQKQKPYAFTVHCVRARRRHWKWAQVTFWCPEQLCHLWLOTLR 120
Db 61 BETDVHGKHGSGKWKQKQKPYAFTVHCVRARRRHWKWAQVTFWCPEQLCHLWLOTLR 120

Qy 121 EMLEKLTSRPKHLLVFINPFGKGQGGKRIYERKVAPLFTLASITTDIIIVTEHANOAKETL 180
Db 121 EMLEKLTSRPKHLLVFINPFGKGQGGKRIYERKVAPLFTLASITTDIIIVTEHANOAKETL 180

Qy 181 YEINIDKYDGIIVCGDGMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPAGS 240
Db 181 YEINIDKYDGIIVCGDGMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPAGS 240

Qy 241 TDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGAGYFGDIIKDS 300
Db 241 TDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHHNSTLLRYSVSLGAGYFGDIIKDS 300

Qy 301 EKRWGLGARYDFSGDKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGFCVCRQSKQ 360
Db 301 EKRWGLGARYDFSGDKTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGFCVCRQSKQ 360

Qy 361 LEEBQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 420
Db 361 LEEBQKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 420

Qy 421 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRKQFTSKHMEDESDSKEGGKRF 480
Db 421 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRKQFTSKHMEDESDSKEGGKRF 480

Qy 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 537
Db 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 537
```

```

RESULT 4
ADN62845
ID   ADN62845 standard; protein; 537 AA.
AC
XX
XX
DT   01-JUL-2004 (first entry)
XX
XX
DE   Human NOV9a.
XX
KW   human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;
KW   anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;
KW   Alzheimer's disease; Parkinson's disease; immune disorder;
KW   haematopoietic disorder; dyslipidaemia; metabolic syndrome X;
KW   wasting disorder.
XX
OS   Homo sapiens.
XX
XX
PN   US2004038223-A1.
XX
PD   26-FEB-2004.
XX
PF   01-OCT-2002; 2002US-00262511.
XX
XX
PR   02-OCT-2001; 2001US-0326483P.
PR   05-OCT-2001; 2001US-0327435P.
PR   05-OCT-2001; 2001US-0327449P.
PR   09-OCT-2001; 2001US-0327917P.
PR   09-OCT-2001; 2001US-0328029P.
PR   09-OCT-2001; 2001US-0328044P.
PR   09-OCT-2001; 2001US-0328056P.
PR   12-OCT-2001; 2001US-0328849P.
PR   15-OCT-2001; 2001US-0329414P.
PR   17-OCT-2001; 2001US-0330142P.
PR   18-OCT-2001; 2001US-0330309P.
PR   22-OCT-2001; 2001US-0341058P.
PR   24-OCT-2001; 2001US-0339266P.
PR   24-OCT-2001; 2001US-0343629P.
PR   29-OCT-2001; 2001US-0349575P.
PR   01-NOV-2001; 2001US-0346357P.
PR   17-APR-2002; 2002US-0373260P.
PR   19-APR-2002; 2002US-0373815P.
PR   19-APR-2002; 2002US-0373817P.
PR   19-APR-2002; 2002US-0373826P.
PR   19-APR-2002; 2002US-0373884P.
PR   22-APR-2002; 2002US-0374977P.
PR   16-MAY-2002; 2002US-0381037P.
PR   16-MAY-2002; 2002US-0381038P.
PR   17-MAY-2002; 2002US-0381042P.
PR   17-MAY-2002; 2002US-0381642P.
PR   28-MAY-2002; 2002US-0383656P.
PR   29-MAY-2002; 2002US-0383831P.
PR   25-JUN-2002; 2002US-0391335P.
XX
XX
(WMIT/) SMITHSON G.
PA (MILL/) MILLET I.
PA (PEYM/) PEYMAN J. A.
PA (KEKU/) KEKUDA R.
PA (JUJJ/) JU J.
PA (LILL/) LI L.
PA (GUOX/) GUO X.
PA (PATT/) PATTURAJAN M.
PA (SPYT/) SPYTEK K. A.
PA (EDIN/) EDINGER S. R.
PA (ELLE/) ELLERMAN K.
PA (MALY/) MALYANKAR U. M.
PA (ORTT/) ORT T.
PA (GORM/) GORMAN L.
PA (ZERH/) ZERHUSEN B. D.
PA (ANDE/) ANDERSON D. W.
PA (ZHON/) ZHONG M.
PA (CATT/) CATTERTON E.

```

---

```

PA (JIWW/) JI W.
PA (MILL/) MILLER C E.
PA (RAST/) RASTELLI L.
PA (STON/) STONE D J.
PA (PENA/) PENA C E A.
PA (SHEN/) SHENOY S G.
PA (SHIM/) SHIMKETS R A.
PA (ROTH/) ROTHENBERG M E.
PA (LEAC/) LEACH M D.
PA (AGEE/) AGEH M L.
PA (BERG/) BERGHS C.
PA (DIPI/) DIPIPPO V A.
PA (EISE/) EISEN A.
PA (GANG/) GANGOLLI E A.
PA (RIEG/) RIEGER D K.
PA (SPAD/) SPADERNA S K.
XX
Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;
PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;
PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;
XX
WPI; 2004-213931/20.
DR N-PSDB; ADN62844.
XX
Isolated NOVX polypeptides and nucleic acids, useful for preventing,
diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.
Claim 1; SEQ ID NO 40; 395pp; English.
XX
The invention relates to isolated NOVX polypeptides and polynucleotides.
CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or
CC treat a medical condition in human related to the aberrant expression and
CC activity of NOVX polypeptides. For example, NOVX polypeptides and
CC polynucleotides may be used to treat disorders associated with decreased
CC expression or activity of NOVX by supplementing the patient our
CC production or to rectify mutations. Conversely, antisense NA molecules
CC may be administered to down regulate expression of NOVX polypeptides by
CC binding with the cells own genes and preventing their expression. NOVX
CC polynucleotides and complementary sequences may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar sequences in samples, and so which patients may be in need of
CC restorative therapy. NOVX polypeptides may also be used as antigens in
CC the production of antibodies and in assays to identify modulators
CC (agonists and antagonists) of the expression and activity of NOVX. The
CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be
CC used to modulate NOVX polynucleotide expression and activity of NOVX
CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as
CC diagnostic agents for detecting the presence of NOVX in samples. NOVX
CC polypeptides and polynucleotides may be used in this way to prevent,
CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious
CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative
CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,
CC haematopoietic disorders, and the various dyslipidaemias, metabolic
CC disturbances associated with obesity, the metabolic syndrome X and
CC wasting disorders associated with chronic diseases and various cancers.
CC They may also be used as antibacterial agents. The present sequence
CC represents the amino acid sequence of a human NOVX protein.
XX
SQ Sequence 537 AA;
Query Match 100.0%; Score 2888; DB 8; Length 537;
Best Local Similarity 100.0%; Pred. No. 2e-274;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 MGATGAAPLQSVLWVKQRCVAVSLPARALLRWRSFGCAGAPGADACSVPSIIV 60
Db 1 MGATGAAPLQSVLWVKQRCVAVSLPARALLRWRSFGCAGAPGADACSVPSIIV 60
Oy 61 EETDVHGKQSGKWKQMKPKYAFTHCVKRRHRHWKVAQVTPWCBEQLCHLWLTLR 120

```

Db 61 BETDVHGKHQSGKQKMEKPYAFTVHCVRARRHRKWAQVTFWCPBEQLCHLWLTQLR 120  
QY 121 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL 180  
Db 121 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL 180  
QY 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDONHPRAVLVPSSLRIGIIPAGS 240  
Db 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDONHPRAVLVPSSLRIGIIPAGS 240  
QY 241 TDCVCYSTVGTSDAETSAHLIIVVGDLSLAMDVSSVHNNSTLLRYSVSLIGYGFYGDIIKDS 300  
Db 241 TDCVCYSTVGTSDAETSAHLIIVVGDLSLAMDVSSVHNNSTLLRYSVSLIGYGFYGDIIKDS 300  
QY 301 EKKRWGLGARYDFSGKLTFLSHHCYEGTVSFPLPAQHTVGSPPDRKPCRCAGFCVCRQSKQ 360  
Db 301 EKKRWGLGARYDFSGKLTFLSHHCYEGTVSFPLPAQHTVGSPPDRKPCRCAGFCVCRQSKQ 360  
QY 361 LEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 420  
Db 361 LEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 420  
QY 421 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRVKXKFQFTSKHMEDESDSKEGGKKRF 480  
Db 421 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRVKXKFQFTSKHMEDESDSKEGGKKRF 480  
QY 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDHS 537  
Db 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDHS 537

RESULT 5

ABB07857  
ID ABB07857 standard; protein; 562 AA.  
AC ABB07857;  
XX  
DT 03-JUL-2002 (first entry)  
DE Human sphingosine kinase-like protein.  
DE  
KW Human sphingosine kinase-like protein; intracellular signalling;  
KW cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;  
KW autoimmune disease; rheumatoid arthritis; Parkinson's disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200228906-A2.  
XX  
PD 11-APR-2002.  
XX  
PF 05-OCT-2001; 2001WO-EP011516.  
XX  
PR 06-OCT-2000; 2000US-0238005P.  
PR 23-AUG-2001; 2001US-0314113P.  
XX  
XX (FARB ) BAYER AG.  
XX  
XX Kossida S, Encinas J;  
PI  
XX  
DR WPI: 2002-340094/37.  
DR N-PSDB; ABL40828.  
XX

PT New reagent for modulating the activity of sphingosine kinase-like  
PT protein polypeptide or polynucleotide and treating cancer, asthma,  
PT allergy, an autoimmune disease, or a central or peripheral nervous system  
PT disorder.  
XX  
PS Claim 25; Fig 11; 120pp; English.

XX The invention relates to a human sphingosine kinase-like protein. The  
CC polypeptide can be expressed by standard recombinant methodology. The  
CC sphingosine kinase-like protein and gene can be used to regulate

CC intracellular signalling and consequently cell proliferation and  
CC apoptosis. Such regulation is useful for treating cancer, allergies (e.g.  
CC asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and  
CC peripheral nervous system disorders (e.g. Parkinson's disease). The  
CC present sequence represents the human sphingosine kinase-like protein  
XX  
SQ Sequence 562 AA;

Query Match 100.0%; Score 2888; DB 5; Length 562;  
Best Local Similarity 100.0%; Pred. No. 2.2e-274;  
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGATGAAPLOSVLVWVQORCAVSLPARALLRWRSPGCGAGAGADACVSPVSEIITAV 60  
Db 26 MGATGAAPLOSVLVWVQORCAVSLPARALLRWRSPGCGAGAGADACVSPVSEIITAV 85  
QY 61 BETDVHGKHQSGKQKMEKPYAFTVHCVRARRHRKWAQVTFWCPBEQLCHLWLTQLR 120  
Db 86 BETDVHGKHQSGKQKMEKPYAFTVHCVRARRHRKWAQVTFWCPBEQLCHLWLTQLR 145  
QY 121 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL 180  
Db 146 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQAKETL 205  
QY 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDONHPRAVLVPSSLRIGIIPAGS 240  
Db 206 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTQRSAGVDONHPRAVLVPSSLRIGIIPAGS 265  
QY 241 TDCVCYSTVGTSDAETSAHLIIVVGDLSLAMDVSSVHNNSTLLRYSVSLIGYGFYGDIIKDS 300  
Db 266 TDCVCYSTVGTSDAETSAHLIIVVGDLSLAMDVSSVHNNSTLLRYSVSLIGYGFYGDIIKDS 325  
QY 301 EKKRWGLGARYDFSGKLTFLSHHCYEGTVSFPLPAQHTVGSPPDRKPCRCAGFCVCRQSKQ 360  
Db 326 EKKRWGLGARYDFSGKLTFLSHHCYEGTVSFPLPAQHTVGSPPDRKPCRCAGFCVCRQSKQ 385  
QY 361 LEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 420  
Db 386 LEEEOKKALYGLEAAEDVEEQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 445  
QY 421 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRVKXKFQFTSKHMEDESDSKEGGKKRF 480  
Db 446 ILIRKCSRFNRLFLIRHTNQDQDFTFVEVYRVKXKFQFTSKHMEDESDSKEGGKKRF 505  
QY 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDHS 537  
Db 506 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDHS 562

RESULT 6

AAM49115  
ID AAM49115 standard; protein; 537 AA.  
XX  
AC AAM49115;  
XX  
DT 20-MAY-2002 (first entry)  
XX  
DE Human ceramide kinase hCERK1.  
XX  
KW Human; ceramide kinase; hCERK1; drug screening; gene therapy;  
KW neurological disease; inflammation; human immunodeficiency virus;  
KW HIV infection; type 2 diabetes; obesity; sepsis; arteriosclerosis;  
KW cancer; neuroprotective; antiinflammatory; anti-HIV; antidiabetic;  
KW anorectic; antibacterial; antiseptic; antiarteriosclerotic; cytostatic;  
KW enzyme.  
XX  
OS Homo sapiens.  
XX  
PN WO200196575-A1.  
XX  
PD 20-DEC-2001.  
XX  
PF 11-JUN-2001; 2001WO-JP004889.

```
XX 14-JUN-2000; 2000JP-00178039.
XX (SANY ) SANKYO CO LTD.
XX Sugiura M, Kono K, Kohama T;
XX WPI; 2002-179513/23.
XX N-PSDB; ABA96945.
XX Human ceramide kinase gene and the enzyme encoded by it for screening
XX substances as drugs for neurological, inflammatory and other disorders.
XX Claim 1; Page 54-57; 61pp; Japanese.
XX This sequence represents a human ceramide kinase designated hCERK1. The
XX invention relates to hCERK1, nucleic acids encoding it, expression
XX vectors and host cells containing hCERK1 nucleic acids, the recombinant
XX production of hCERK1 and antibodies specific for hCERK1. The invention
XX also encompasses methods of isolating hCERK1 from samples, the use of
XX hCERK1 in drug screening, and the use of hCERK1 nucleic acid sequences in
XX gene therapy. hCERK1 mediates the ATP-dependent 1-phosphorylation of
XX ceramides and can be used to screen for therapeutic and preventive agents
XX for a wide range of disorders. Such disorders include neurological
XX disease, inflammation, human immunodeficiency virus (HIV) infection, type
XX 2 diabetes, obesity, sepsis, arteriosclerosis and cancer
XX Sequence 537 AA;
XX
Query Match 99.7%; Score 2880; DB 5; Length 537;
Best Local Similarity 99.6%; Pred. No. 1.2e-273;
Matches 535; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MGATGAEPLOSVLWVKQRCVAVSLEPARALLRWSPGCGAGAPGADACVPVSEIIAV 60
Db 1 MGATGAEPLOSVLWVKQRCVAVSLEPARALLRWSPGCGAGAPGADACVPVSEIIAV 60
Qy 61 EETDVHGKHGQSGKWQMEKPYAFTVHCVKARRHRWKQAQVTFWCPEEQCHLWLQTLR 120
Db 61 EETDVHGKHGQSGKWQMEKPYAFTVHCVKARRHRWKQAQVTFWCPEEQCHLWLQTLR 120
Qy 121 EMLEKLTSRPKHLLVFINPFGKGQGRKIYERKVAPLFTLASITTDIIIVTEHANQA 180
Db 121 EMLEKLTSRPKHLLVFINPFGKGQGRKIYERKVAPLFTLASITTDIIIVTEHANQA 180
Qy 181 YEINIDKYDGIYCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGI 240
Db 181 YEINIDKYDGIYCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGI 240
Qy 241 TDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHNSTLLRYSVSLGFGYGDIIKDS 300
Db 241 TDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHNSTLLRYSVSLGFGYGDIIKDS 300
Qy 301 EKKRWGLGARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRAGCFVCRQ 360
Db 301 EKKRWGLGARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRAGCFVCRQ 360
Qy 361 LEEQKALYGLEAAEDVEEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
Db 361 LEEQKALYGLEAAEDVEEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420
Qy 421 ILIRKCSRFNRLFLIRHTNQDQDFTFVVEYVRVKKQFTSKHMEDESDLKEGGKRF 480
Db 421 ILIRKCSRFNRLFLIRHTNQDQDFTFVVEYVRVKKQFTSKHMEDESDLKEGGKRF 480
Qy 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDHS 537
Db 481 GHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDHS 537
RESULT 7
ABR56302
ID ABR56302 standard; protein; 481 AA.
```

```
XX ABR56302;
XX
XX 20-NOV-2003 (first entry)
XX Human Sphingosine kinase 4-related protein.
XX Human; enzyme; haemostatic; sphingosine kinase 4; SPHK4;
XX platelet transfusion; platelet stabiliser.
XX Homo sapiens.
XX WO2003031627-A1.
XX
XX 17-APR-2003.
XX
XX 28-SEP-2001; 2001WO-JP008537.
XX
XX 28-SEP-2001; 2001WO-JP008537.
XX (HOKK-) HOKKAIDO TECHNOLOGY LICENSING OFFICE CO.
XX (CHBI-) CHEM BIOLOGY INST.
XX Igarashi Y, Kihara A;
XX WPI; 2003-354917/33.
XX
XX Platelet derived polypeptides with sphingosine kinase activity for
XX treatment of sphingosine related disorders.
XX
XX Example 4; Fig 4; 39pp; Japanese.
XX
XX The present invention relates to human sphingosine kinase 4 (SPHK4;
XX ABR56301). The kinase can be used for the diagnosis and treatment of
XX sphingosine related disorders. The kinase can also be potentially used
XX for controlling toxicity of platelet transfusion and as a platelet
XX stabiliser. The present sequence was used to illustrate the invention
XX
XX Sequence 481 AA;
XX
Query Match 89.6%; Score 2588; DB 6; Length 481;
Best Local Similarity 100.0%; Pred. No. 5.4e-245;
Matches 481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 57 IIAVEETDVHGKHGQSGKWQMEKPYAFTVHCVKARRHRWKQAQVTFWCPEEQCHLWL 116
Db 1 IIAVEETDVHGKHGQSGKWQMEKPYAFTVHCVKARRHRWKQAQVTFWCPEEQCHLWL 60
Qy 117 QTLREMLEKLTSRPKHLLVFINPFGKGQGRKIYERKVAPLFTLASITTDIIIVTEHANQA 176
Db 61 QTLREMLEKLTSRPKHLLVFINPFGKGQGRKIYERKVAPLFTLASITTDIIIVTEHANQA 120
Qy 177 KETLYEINIDKYDGIYCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGI 236
Db 121 KETLYEINIDKYDGIYCVGGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLPVSSLRIGI 180
Qy 237 PAGSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHNSTLLRYSVSLGFGYGDII 296
Db 181 PAGSTDCVCYSTVGTSDAETSALHIVVGDLSAMDVSSVHNSTLLRYSVSLGFGYGDII 240
Qy 297 IKDSEKKRWGLGARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRAGCFVCRQ 356
Db 241 IKDSEKKRWGLGARYDFSGLKTFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRAGCFVCRQ 300
Qy 357 SKQOLEEQKALYGLEAAEDVEEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDG 416
Db 301 SKQOLEEQKALYGLEAAEDVEEQVVCCKFLAINATNMSCACRRSPRGLSPAHLGDG 360
Qy 417 SSDLLILIRKCSRFNRLFLIRHTNQDQDFTFVVEYVRVKKQFTSKHMEDESDLKEGG 476
Db 361 SSDLLILIRKCSRFNRLFLIRHTNQDQDFTFVVEYVRVKKQFTSKHMEDESDLKEGG 420
Qy 477 KGRFGHICSSHPSCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDHS 536
```

```

Db 421 KRRFGHICSSHPSCCTVSNWNCDEVLHSPAIEVRVHCOLVRLFARGIEENKPDUSH 480
Qy 537 S 537
Db 481 S 481

RESULT 8
AA96059
ID AA96059 standard; protein; 460 AA.
AC AA96059;
XX
DT 05-DEC-2000 (first entry)
DE Human sphingosine kinase C.
KW Sphingosine kinase C; SKC; human; drug screening; infection;
KW antiinflammatory; antiallergic; anticancer; inflammation; allergy;
KW cancer; therapy; diagnosis.
XX
OS Homo sapiens.
XX
PN WO200052173-A2.
XX
PD 08-SEP-2000.
XX
PF 02-MAR-2000; 2000WO-CA000223.
XX
PR 02-MAR-1999; 99US-0122516P.
XX
PA (ALLX ) NPS ALLELIX CORP.
XX
PI Munroe D, Gupta A, Falzone GR;
XX
DR WPI; 2000-572185/53.
DR N-PSDB; AAA50510.
XX
New human sphingosine kinase A, B and C polynucleotides and polypeptides
useful in e.g. chromosome and gene mapping, and detecting inflammation or
disease associated with abnormal levels of sphingosine kinase expression.
XX
PS Disclosure; Fig 9; 81pp; English.
XX
XX
The present sequence is that of human sphingosine kinase C (SKC), an
enzyme that phosphorylates sphingosine to form sphingosine 1-phosphate.
The sequence was deduced from that of a polynucleotide (see AAA50510)
isolated from an HeLa cDNA library. The invention provides
polynucleotides (see AAA50508-10) and polypeptides (see AA96057-59) for
the human sphingosine kinase (SK) homologues SKA, SKB and SKC. The
polypeptides can be obtained using recombinant DNA methods, and host
cells containing expression vectors including SK polynucleotides are used
in a claimed method of screening for compounds that inhibit or activate
human SK activity. Human SK specific antibodies, inhibitors, ligands or
their analogues can be used as bioactive agents to treat inflammation or
disease including viral, bacterial or fungal infections, allergic
responses, mechanical injury associated with trauma, hereditary diseases,
lymphoma or carcinoma, and other conditions with activate the genes of
kidney, lung, heart, lymphoid or tissues of the nervous system
XX
SQ Sequence 460 AA;
Query Match 85.3%; Score 2463; DB 3; Length 460;
Best Local Similarity 99.6%; Pred. No. 9, 9e-233;
Matches 458; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 78 MEKPYAFTVHCYKARRHRKWAQVTFWCPEQLCHLWQLTREMLEKLTSPKHLVFI 137
Db 1 MEKPYAFTVHCYKARRHRKWAQVTFWCPEQLCHLWQLTREMLEKLTSPKHLVFI 60
Qy 138 NPFSGKGGKRIYERKVAFLTLASITTDIIVTEHANOAKETLYEINIDKDYGVCGD 197

```

```

Db 61 NPFSGKGGKRIYERKVAFLTLASITTDIIVTEHANOAKETLYEINIDKDYGVCGD 120
Qy 198 GMPSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGSTDVCYSTVGTSDAETS 257
Db 121 GMPSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGSTDVCYSTVGTSDAETS 180
Qy 258 ALHIWVGDSLAMDYSSVHHNSTLLRYSVSLIGYFYGDIIKDSEKKRWGLARYDFSLK 317
Db 181 ALHIWVGDSLAMDYSSVHHNSTLLRYSVSLIGYFYGDIIKDSEKKRWGLARYDFSLK 240
Qy 318 TFLSHHCVEGTVSFLPAQHTVGSPPDRKPCRAGCFVCRQSKQOLEEEOKKALYGLEAAED 377
Db 241 TFLSHHCVEGTVSFLPAQHTVGSPPDRKPCRAGCFVCRQSKQOLEEEOKKALYGLEAAED 300
Qy 378 VEEQVVCGRFLAINATNMSCACRRSPRGLSPAHLGDGSSDLILIRKCSRNFNLFRLIR 437
Db 301 VEEQVVCGRFLAINATNMSCACRRSPRGLSPAHLGDGSSDLILIRKCSRNFNLFRLIR 360
Qy 438 HTNQDQDFTFVEVYRVKXKQFTSKMEDSDSLKEGKKRFGHICSSHPSCCCTVSN 497
Db 361 HTNQDQDFTFVEVYRVKXKQFTSKMEDSDSLKEGKKRFGHICSSHPSCCCTVSN 420
Qy 498 SWNCDGEVLHSPAIEVRVHCOLVRLFARGIEENKPDUSH 537
Db 421 SWNCDGEVLHSPAIEVRVHCOLVRLFARGIEENKPDUSH 460

RESULT 9
AAE07884
ID AAE07884 standard; protein; 471 AA.
AC AAE07884;
XX
DT 01-NOV-2001 (first entry)
DE Human sphingosine kinase (SphK) protein #2.
XX
KW Human; sphingosine kinase; SphK; restenosis; ischaemia; gene therapy;
KW antisense therapy; cancer; sphingolipid; signalling molecule; apoptosis;
KW cytosstatic; tumour necrosis factor-alpha; TNF; atherosclerosis; lymphoma;
KW leukaemia; vasotropic; cell proliferative disorder; vascular disease.
XX
OS Homo sapiens.
XX
PN WO200160990-A2.
XX
PD 23-AUG-2001.
XX
PF 14-FEB-2001; 2001WO-US004789.
XX
PR 14-FEB-2000; 2000US-0182360P.
XX
PR 22-MAR-2000; 2000US-0191261P.
XX
PA (CURA-) CURAGEN CORP.
PA (GETH ) GENENTECH INC.
XX
Rastelli L;
XX
WPI; 2001-514770/56.
XX N-PSDB; AAD14426.
XX
An isolated Sphingosine kinase polypeptide useful for treating a SphK-
associated disorder especially cancer, restenosis or ischemia in a human.
XX
Claim 1; Fig 1; 107pp; English.
XX
The present invention relates to sphingosine kinase (SphK) polypeptides
and nucleic acids encoding them. SphK is useful for treating a SphK-
associated disorder especially cancers such as leukaemia, lymphoma,
ovarian, breast, lung, colon, testicular, stomach and skin,
atherosclerosis, restenosis or ischaemia and cell proliferative disease
or disorder associated with vascular diseases. SphK gene is used in gene
therapy and antisense-therapy. Sphingolipids serving as signalling

```



CC	molecules, have recently emerged as regulators of cell growth,
CC	differentiation, diverse cell phenotypes and cell death. Activation of
CC	SphK by tumour necrosis factor (TNF)-alpha inhibits apoptosis in human
CC	endothelial cells. The present sequence is human sphingosine kinase
CC	(SphK) protein
XX	
QQ	Sequence 471 AA;
Query Match	85.1%; Score 2456.5; DB 4; Length 471;
Best Local Similarity	97.5%; Pred. No. 4.4e-232;
Matches 459; Conservative	1; Mismatches 0; Indels 11; Gaps 1;
QY	78 MEKPYATVTCVKARHRKWAQVTFCPEEQI <sup>CHLW</sup> LTREML <sup>EKLTSR</sup> PKHLIVFI 137
DB	1 MEKPYATVTCVKARHRKWAQVTFCPEEQI <sup>CHLW</sup> LTREML <sup>EKLTSR</sup> PKHLIVFI 60
QY	138 NPGGKGOGKRIYERKVAPLFTLASITTDII-----VTEHANQAETLYEINID 186
DB	61 NPGGKGOGKRIYERKVAPLFTLASITTDIIIGNKFYNYVEVITEHANQAETLYEINID 120
QY	187 KYDGIVCVGDMSEVLHGLIGRTQSAGVDQNHPRAVLVPSSLIRIGIIPAGSTDVCVY 246
DB	121 KYDGIVCVGDMSEVLHGLIGRTQSAGVDQNHPRAVLVPSSLIRIGIIPAGSTDVCVY 180
QY	247 STVGTSDAETSALHIVVGDLSAMDVSVHHNSTLLRYSVSLGFGFYGDIIKDSEKKRWL 306
DB	181 STVGTSDAETSALHIVVGDLSAMDVSVHHNSTLLRYSVSLGFGFYGDIIKDSEKKRWL 240
QY	307 GLARYDFSGLKTFLSHCYEGTWSFLPAOHTVGSPRDKPCRACGFVCRSKOOLEBEOK 366
DB	241 GLARYDFSGLKTFLSHCYEGTWSFLPAOHTVGSFRDKPCRACGFVCRSKOOLEBEOK 300
QY	367 KALYGLEAAEDVEBQVVCVKFLAINATNNSCACRRSPRGISLPAAHLDGSSDLILIRKC 426
DB	301 KALYGLEAAEDVEBQVVCVKFLAINATNNSCACRRSPRGISLPAAHLDGSSDLILIRKC 360
QY	427 SRNFLRLIRHTNQODQFDFTFEVYRVKKFQFTSKHMEDESDLKEGGKRFGHTCSS 486
DB	361 SRNFLRLIRHTNQODQFDFTFEVYRVKKFQFTSKHMEDESDLKEGGKRFGHTCSS 420
QY	487 HPSCCCTVSNSWNCDGEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSSH 537
DB	421 HPSCCCTVSNSWNCDGEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSSH 471
RESULT 10	
ADP55248	
ID	ADP55248 standard; protein; 531 AA.
XX	
AC	ADP55248;
XX	
DT	18-NOV-2004 (first entry)
XX	
DE	Human PRO protein sequence SEQ ID NO:1224.
XX	
KW	human; PRO; immune related disease; inflammatory immune response;
KW	immune response stimulation; antiallergic; antianaemic; antiarthritic;
KW	antiaethmatic; antidiabetic; antiinflammatory; antipsoriatic;
KW	antipneumatic; antithyroid; CNS; dermatological; gastrointestinal;
KW	haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;
KW	nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;
KW	virucide; gene therapy.
OS	Homo sapiens.
XX	
FN	WO2004039956-A2.
XX	
PD	13-MAY-2004.
XX	
PF	28-OCT-2003; 2003WO-US034381.
XX	
PR	29-OCT-2002; 2002US-0422472P.
XX	

QY 361 LEEOKKALYGLEAAEDVEEVOVCGKFLAINATNMSCACRSPGSLSPAHLGDGSSDL 420  
DB 361 LEEEEKKALYGLENAEEVEEWOVTCGKFLAINATNMSCACRSPGSLSPAHLGDGSSDL 420  
QY 421 ILIRKCSRNFNRLIRHTNQDQDFTFVEVYRVKKFOFTSKHMEDESDLKEGGKRF 480  
DB 421 ILIRKCSRNFNRLIRHTNQDQDFTFVEVYRVKKFHTSKHVEDESDNSKEQEKORF 480  
QY 481 GHICSHSPCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEE 529  
DB 481 GKICKDRPSCTSASRSSWNCDEVMHSPAIEVRVHCOLVRLFARGIEE 529

## RESULT 11

ADSI11054  
ID ADSI11054 standard; protein; 536 AA.

AC ADSI11054;

DT 16-DEC-2004 (first entry)

DE Human therapeutic protein - SEQ ID 1291.

KW antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnerary;  
KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;  
KW aplastic anaemia; cancer; wound healing; gene therapy.

OS Homo sapiens.

FN WO2004080148-A2.

PD 23-SEP-2004.

PF 30-SEP-2003; 2003WO-US030720.

PR 02-OCT-2002; 2002US-0416186P.

PA (NUVE-) NUVELO INC.

PI Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y;  
PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;

XX WPI; 2004-668857/65.

DR N-PSDB; ADS10370.

PT New polynucleotide, useful in preparing a composition for diagnosing or  
PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,  
PT aplastic anemia or cancer for promoting wound healing.

PS Claim 20; SEQ ID NO 1291; 718pp; English.

XX The invention relates to a novel isolated polynucleotide and the encoded  
CC polypeptide. The molecules of the invention demonstrate antiinflammatory,  
CC neuroprotective, antianaemic, cytostatic and vulnerary activities and may  
CC be useful in preparing a composition for diagnosing or treating  
CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell  
CC disorders, such as aplastic anaemia or cancer, as well as for promoting  
CC wound healing. The molecules may also be utilised during gene therapy  
CC procedures. The current sequence is that of a human therapeutic protein  
CC of the invention. The current sequence is not shown explicitly within the  
CC specification but can be accessed from the WIPO web-site.

SQ Sequence 536 AA;

Query Match 81.4%; Score 2350; DB 8; Length 536;

Best Local Similarity 78.1%; Pred. No. 1.6e-221;

Matches 452; Conservative 2; Mismatches 9; Indels 116; Gaps 4;

QY 32 LRWWR-----SPGPGAGAPGADACSPVPSIIAEEVDVHGKHGSGKWQMKPYAFT 85

DB 1 MEWGRVFCVEMPAGRGV-FTDACSPVPSIIAEEVDVHGKHGSGKWQMKPYAFT 59

QY 86 VHCVRARRHRKWAQVTFWCPEQLCHLWLTQTLREMLEKLTSPKILLVFINPFGKGXQ 145

DB 60 VHCVRARRHRKWAQVTFWCPEQLCHLWLTQTLREMLEKLT----- 100  
QY 146 GKRIYERKVAFLFTLASITTDIIITEHANOAKETLYEINIDKYDGVCGVGGDGFSEVLH 205  
DB 101 -----ITEHANOAKETLYEINIDKYDGVCGVGGDGFSEVLH 137  
QY 206 GLIGRTQRSAGVDQNHPRAVLVPSSLRTIGIIPA----- 238  
DB 138 GLIGRTQRSAGVDQNHPRAVLVPSSLRTIGIIPAGFANDLAGDEVSVLSGPVSGPBGGMV 197  
QY 239 -----GSTDCVCYSTVGTSDAETSA 258  
DB 198 HTFTLVLTALGCEHRSHPHFMDERTGTGHHVGFGENAGGLDGTDCVCYSTVGTSDAETSA 257  
QY 259 LHIIVVGDSLAMDVSVVHNSLTLRYSVSLGYPGYDIIKDSEKKRWLGLARYDFSGLKT 318  
DB 258 LHIIVVGDSLAMDVSVVHNSLTLRYSVSLGYPGYDIIKDSEKKRWLGLARYDFSGLKT 317  
QY 319 FLSHHCYEGTVSFLPAQHTVGSPRDRKPCFVCRQSKQLEEEQKKALYGLEAADV 378  
DB 318 FLSHHCYEGTVSFLPAQHTVGSPRDRKPCFVCRQSKQLEEEQKKALYGLEAADV 377  
QY 379 BEWQVCGKFLAINATNMSCACRSPGSLSPAHLGDGSSDLILIRKCSRNFNRLIRH 438  
DB 378 BEWQVCGKFLAINATNMSCACRSPGSLSPAHLGDGSSDLILIRKCSRNFNRLIRH 437  
QY 439 TNQDQDFTFVEVYRVKKFOFTSKHMEDESDLKEGGKRFHGTCSHSPCCCTVSNSS 498  
DB 438 TNQDQDFTFVEVYRVKKFOFTSKHMEDESDLKEGGKRFHGTCSHSPCCCTVSNSS 497  
QY 499 WNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 537  
DB 498 WNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSHS 536

## RESULT 12

ABR56301

ID ABR56301 standard; protein; 416 AA.

AC ABR56301;

DT 20-NOV-2003 (first entry)

DE Human Sphingosine kinase 4.

KW Human; enzyme; haemostatic; sphingosine kinase 4; SPHK4;  
KW platelet transfusion; platelet stabiliser.

OS Homo sapiens.

PN WO2003031627-A1.

PD 17-APR-2003.

PF 28-SEP-2001; 2001WO-JP008537.

PR 28-SEP-2001; 2001WO-JP008537.

PA (HOKK-) HOKKAIDO TECHNOLOGY LICENSING OFFICE CO.

PI (CHBI-) CHEM BIOLOGY INST.

PI Igaraeahi Y, Kihara A;

XX WPI; 2003-354917/33.

DR N-PSDB; ACC70838.

PT Platelet derived polypeptides with sphingosine kinase activity for  
PT treatment of sphingosine related disorders.

PS Claim 2; Page 29-30; 39pp; Japanese.

CC The present sequence is the protein sequence for human sphingosine kinase

CC 4 (SPHK4). The kinase can be used for the diagnosis and treatment of  
CC sphingosine related disorders. The kinase can also be potentially used  
CC for controlling toxicity of platelet transfusion and as a platelet  
CC stabiliser

XX SQ Sequence 416 AA;

Query Match 76.5%; Score 2210; DB 6; Length 416;  
Best Local Similarity 100.0%; Pred. No. 6.4e-208;  
Matches 416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 122 MLEKLSRPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLY 181  
Dy 1 MLEKLSRPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLY 60  
Qy 182 EINDKVDGIVCVGDDGFMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPAGST 241  
Dy 61 EINDKVDGIVCVGDDGFMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPAGST 120  
Qy 242 DCVCYSTVGTSDAETLSALHIVVGDLSAMDVSSVHNSTLLRYSVLLGYGFGYDIIKDSE 301  
Dy 121 DCVCYSTVGTSDAETLSALHIVVGDLSAMDVSSVHNSTLLRYSVLLGYGFGYDIIKDSE 180  
Qy 302 KRWGLGLARYDFSLKTFSLHCHCEYGVTSFLPAQHTVGSPPDRKPCRAGCFVCRSQKQL 361  
Dy 181 KRWGLGLARYDFSLKTFSLHCHCEYGVTSFLPAQHTVGSPPDRKPCRAGCFVCRSQKQL 240  
Qy 362 EEOKKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSGSDLI 421  
Dy 241 EEOKKALYGLEAAEDVEEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSGSDLI 300  
Qy 422 LIRKCSRFNRLIRHTNQDQDFTFVEYVRVKKQFTSKHMEDESDLKEGGKRRFG 481  
Dy 301 LIRKCSRFNRLIRHTNQDQDFTFVEYVRVKKQFTSKHMEDESDLKEGGKRRFG 360  
Qy 482 HICSHSPSCCTTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIBENPKPDSHS 537  
Dy 361 HICSHSPSCCTTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIBENPKPDSHS 416

RESULT 13

ABGI3541  
ID ABGI3541 standard; protein; 746 AA.

XX AC ABGI3541;

DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #13532.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS77728.

XX PT New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

XX PS Claim 20; SEQ ID NO 43900; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
CC patent did not appear in the printed specification. Note: The sequence data for this  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 746 AA;

Query Match 76.5%; Score 2208.5; DB 4; Length 746;  
Best Local Similarity 75.8%; Pred. No. 2.1e-207;  
Matches 430; Conservative 6; Mismatches 16; Indels 115; Gaps 6;

Qy 32 LRWR-----SCPAGAGACADACVPVSEIIAIVETDVHGKHQSGKQKMEKPYAFT 85  
Dy 200 MEWRRRVFCVEMFAEGGV-FTDACSVPVSEIIAIVETDVHGKHQSGKQKMEKPYAFT 258  
Qy 86 VHCVKARRHRHWKWAQVTFWCPEEQCLHMLQTLREMLEKLTSPKHLVFINPFGKGQ 145  
Dy 259 VHCVKARRHRHWKWAQVTFWCPEEQCLHMLQTLREMLEKLTSPKHLVFINPFGKGQ 299  
Qy 146 GKRIYERKVAFLFTLASITTDIIIVTEHANOAKETLYEINIDKYDGIIVCVGDDGFMFSEVLH 205  
Dy 300 -----ITEHANOAKETLYEINIDKYDGIIVCVGDDGFMFSEVLH 336  
Qy 206 GLIGRTORSAGVDQNHPRVLPSSLRIGIIPA-----GSTDCVCYSTVGTGTS 252  
Dy 337 GLIGRTORSAGVDQNHPRVLPSSLRIGIIPAHHVGPGENAGGLDSTDCVCYSTVGTGTS 396  
Qy 253 DAETLSALHIVV-----G 264  
Dy 397 DAETLSALHIVVCCPEARKPPASRHTACGSGHGLCLGCOLWNACLEASRLQSRMQSPG 456  
Qy 265 DSLAMDVSSVHNSTLLRYSVLLGYGFGYDIIKDESKRWGLGLARYDFSLKTFSLHCH 324  
Dy 457 DSLAMDVSSVHNSTLLRYSVLLGYGFGYDIIKDESKRWGLGLARYDFSGIKTFSLHCH 516  
Qy 325 YEGTVSFLPAQHTVGSPPDRKPCRAGCFVCRSQKQLEEBQKALYGLEAAEDVEEQVV 384  
Dy 517 YEGTVSFLPAQHTVGSPPDRKPCRAGCFVCRSQKQLEEBQKALYGLEAAEDVEEQVV 576  
Qy 385 CGKFLAINATNMSCACRRSPRGLSPAHLGDSGSDLIIRKCSRFNRLIRHTNQDQ 444  
Dy 577 CGKFLAINATNMSCACRRSPRGLSPAHLGDSGSDLIIRKCSRFNRLIRHTNQDQ 636  
Qy 445 FDFTFVEYVRVKKQFTSKHMEDESDLKEGGKRRFGIHCSSHPSCCTTVSNSSWNCDE 504  
Dy 637 FDFTFVEYVRVKKQFTSKHMEDESDLKEGGKRRFGIHCSSHPSCCTTVSNSSWNCDE 696  
Qy 505 VLHSPAIEVRVHCOLVRLFARGIBENP 531  
Dy 697 VLHSPAIEVR-----AVVMARGCKVP 718

```
RESULT 14
ABG13543
ID ABG13543 standard; protein; 727 AA.
XX AC ABG13543;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #13534.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX XX (HYSE-) HYSEQ INC.
XX PA Drmanac RT, Liu C, Tang YT;
XX PI WPI; 2001-639362/73.
XX DR N-PSDB; AAS77730.
XX XX
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 20; SEQ ID NO 43902; 103pp; English.
XX XX
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activity. The
XX CC polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG00010-ABG030377 represent novel human diagnostic
XX CC amino acid sequences of the invention. Note: The sequence data for this
XX CC patent did not appear in the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 727 AA;
Query Match 59.4%; Score 1714.5; DB 4; Length 727;
Best Local Similarity 63.7%; Pred No. 7,4e-159;
Matches 345; Conservative 0; Mismatches 0; Indels 197; Gaps 3;
QY 170 TEHANOAKETLYEINIDKYDG----- 190
DB 1 TEHANOAKETLYEINIDKYDGCTAHSRMPCCPRNHPAGKAAPVQGVGVGGQRAFPWEP 60
QY 191 -----IVCVGGDGMFSEVLHGL 207
|||||
```

```
Db 61 GPKOMPAGQVCRRLPLLLKLEASGLLRSEGRTRCSPDRLCSSCSIVCVGGDGMFSEVLHGL 120
QY IGRTORAGVDQNHPRAVLVPSSLRIGIIP----- 238
|||
Db 121 IGRTORAGVDQNHPRAVLVPSSLRIGIIPAGPANDLAGDEVSVLSGVPSPGEGGMVHT 180
QY -----GSTDCVCYSTVGTSDARTSALH 260
|||
Db 181 FTLVTALGCEHRSHPHFMDERRTGCEHVPGENAGGLDSTDCVCYSTVGTSDARTSALH 240
QY 261 IIV-----GDSLAMDVS 272
|||
Db 241 IVGCCPEARKPPASRHTACSGHGQLCLGQLNACLCEASRLQSRMQSPGDSLAMDVS 300
QY 273 SVHENSTLLRYSVSLGFGFYGDIIKDKSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFL 332
|||
Db 301 SVHENSTLLRYSVSLGFGFYGDIIKDKSEKKRWGLARYDFSGLKTFLSHHCYEGTVSFL 360
QY 333 PAQHTVGSPPDRKPCRAGCFVCRQSKQQLBEQKALYGLEAAEDVEWQVVCVKFLAIN 392
|||
Db 361 PAQHTVGSPPDRKPCRAGCFVCRQSKQQLBEQKALYGLEAAEDVEWQVVCVKFLAIN 420
QY 393 ATNMSCACRRSPRGLSPAHLGDGSSDLILIRKCSRNFNLFRLIRHTNQDQDFTFVEV 452
|||
Db 421 ATNMSCACRRSPRGLSPAHLGDGSSDLILIRKCSRNFNLFRLIRHTNQDQDFTFVEV 480
QY 453 YRVKKFQFTSKHMEDESDLEKGGKQKFGHICSSHPSCCCTVSNSSMNCDEVLHSPAIE 512
|||
Db 481 YRVKKFQFTSKHMEDESDLEKGGKQKFGHICSSHPSCCCTVSNSSMNCDEVLHSPAIE 540
QY 513 VR 514
||
Db 541 VR 542
RESULT 15
ABG07854
ID ABG07854 standard; protein; 326 AA.
XX AC ABG07854;
XX DT 03-JUL-2002 (first entry)
XX DE Human sphingosine kinase-like protein.
XX KW Human sphingosine kinase-like protein; intracellular signalling;
XX KW cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;
XX KW autoimmune disease; rheumatoid arthritis; Parkinson's disease.
XX OS Homo sapiens.
XX PN WO200228906-A2.
XX PD 11-APR-2002.
XX PF 05-OCT-2001; 2001WO-EP011516.
XX PR 06-OCT-2000; 2000US-0238005P.
XX PR 23-AUG-2001; 2001US-0314113P.
XX PA (FARB ) BAYER AG.
XX PI Kossida S, Encinas J;
XX DR WPI; 2002-340094/37.
XX DR N-PSDB; ABL40822.
XX PT New reagent for modulating the activity of sphingosine kinase-like
XX PT protein polypeptide or polynucleotide and treating cancer, asthma,
XX PT allergy, an autoimmune disease, or a central or peripheral nervous system
XX PT disorder.
XX PS Claim 25; Fig 2; 120pp; English.
```

XX The invention relates to a human sphingosine kinase-like protein. The  
CC polypeptide can be expressed by standard recombinant methodology. The  
CC sphingosine kinase-like protein and gene can be used to regulate  
CC intracellular signalling and consequently cell proliferation and  
CC apoptosis. Such regulation is useful for treating cancer, allergies (e.g.  
CC asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and  
CC peripheral nervous system disorders (e.g. Parkinson's disease). The  
CC present sequence represents the human sphingosine kinase-like protein  
XX  
SQ Sequence 326 AA;  
  
Query Match 56.8%; Score 1640.5; DB 5; Length 326;  
Best Local Similarity 96.3%; Pred. No. 4.4e-152;  
Matches 314; Conservative 1; Mismatches 0; Indels 11; Gaps 1;  
  
QY 130 PKHLLVFNPFGGKQGKRIYERKVPAPLFTLASITTDII-----VTEHANQAKE 178  
DB 1 PKHLLVFNPFGGKQGKRIYERKVPAPLFTLASITTDIIIGNKFYVNVYEVITEHANQAKE 60  
  
QY 179 TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPA 238  
DB 61 TLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPA 120  
  
QY 239 GSTDCVCYSTVGTSDAETSA LHI VVGDSLAMDVSSVHHNSTLLRYSVSLLGYGFYGDIIK 298  
DB 121 GSTDCVCYSTVGTSDAETSA LHI VVGDSLAMDVSSVHHNSTLLRYSVSLLGYGFYGDIIK 180  
  
QY 299 DSEKKRWLGLARYDFSGJKTF LSHHCYEGTVSF LPAQHTVGSPRDRKPCRAGCFVCRQSK 358  
DB 181 DSEKKRWLGLARYDFSGJKTF LSHHCYEGTVSF LPAQHTVGSPRDRKPCRAGCFVCRQSK 240  
  
QY 359 QOLEEQQKALYGLEAAEDVEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 418  
DB 241 QOLEEQQKALYGLEAAEDVEWQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSS 300  
  
QY 419 DLILIRKCSRFNFLRLIRHTNQDDQ 444  
DB 301 DLILIRKCSRFNFLRLIRHTNQDDQ 326

Search completed: September 3, 2005, 04:37:40  
Job time : 111.661 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 6, 2005, 02:35:14 ; Search time 4294.05 Seconds  
(without alignments)  
3678.672 Million cell updates/sec

Title: US-10-631-958-2

Perfect score: 1717

Sequence: 1 PKLLVFNPFGKGGKGRK1.....KCSRPNFLRLIRHTNQDDQ 326

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO.spool/US10631958/runat.02092005.165815.4533/app.query.fasta\_1.1941  
-DB=GenEmbl -QWTF=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOFCU=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10631958 @CGN 1.1.11756 @runat.02092005.165815.4533 -NCPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1717	100.0	979	6	AX456998 Sequence
2	1717	100.0	1840	6	AX224383 Sequence
3	1640.5	95.5	1654	9	CR456404 Homo sapi
4	1640.5	95.5	2042	9	AJ457828 Homo sapi

5	1640.5	95.5	4171	6	BD183468	BD183468	Novel gen
6	1640.5	95.5	4171	9	AB051433	AB051433	Homo sapi
7	1640.5	95.5	4413	6	AX457006	AX457006	Sequence
8	1640.5	95.5	4445	9	AB079066	AB079066	Homo sapi
9	1633.5	95.1	4432	6	AR541900	AR541900	Sequence
10	1632.5	95.1	4463	6	BD102675	BD102675	Ceramide
11	1503.5	87.6	2830	10	AB079067	AB079067	Mus muscu
12	1503.5	87.6	3661	10	AK129416	AK129416	Mus muscu
13	1454	84.7	1459	6	CQ730476	CQ730476	Sequence
14	1452	84.6	1772	9	BC067255	BC067255	Homo sapi
15	1251.5	72.9	1450	5	CR386590	CR386590	Gallus ga
16	1140	66.4	1520	5	BC074110	BC074110	Xenopus l
17	1117.5	65.1	2494	5	BC074350	BC074350	Xenopus l
18	811	47.2	474	6	AX457001	AX457001	Sequence
19	650.5	37.9	2333	3	AK112750	AK112750	Ciona int
20	613	35.7	1810	3	AY061001	AY061001	Drosophil
21	613	35.7	2064	6	AR509346	AR509346	Sequence
22	613	35.7	2173	6	CQ608040	CQ608040	Sequence
23	613	35.7	3178	3	BT014925	BT014925	Drosophil
24	613	35.7	10337	6	CQ608039	CQ608039	Sequence
25	613	35.7	22353	2	AC017269	AC017269	Drosophil
26	613	35.7	181009	3	AC007575	AC007575	Drosophil
27	613	35.7	295289	3	AE003603	AE003603	Drosophil
28	566	33.0	1084	6	AR510195	AR510195	Sequence
29	508.5	29.6	329	6	AX457002	AX457002	Sequence
30	498.5	29.0	2257	8	AY362552	AY362552	Arabidops
31	473	27.5	2263	8	AY035126	AY035126	Arabidops
32	436	25.4	1599	9	AY357073	AY357073	Homo sapi
33	414.5	24.1	564	6	AX718996	AX718996	Sequence
34	389	22.7	1458	8	BT006621	BT006621	Arabidops
35	389	22.7	1757	8	AY128394	AY128394	Arabidops
36	389	22.7	70890	9	HSDAS9H18	AL096766	Human DNA
37	386	22.5	181772	2	AC007689	AC007689	Homo sapi
38	356.5	20.8	219077	2	AC116764	AC116764	Mus muscu
39	355	20.7	1799	8	AK072068	AK072068	Oryza sat
40	355	20.7	1817	8	AK107062	AK107062	Oryza sat
41	347	20.2	240483	2	AC095955	AC095955	Rattus no
42	347	20.2	260331	2	AC135400	AC135400	Rattus no
43	339.5	19.8	2451	9	AK129976	AK129976	Homo sapi
44	329	19.2	79048	2	AC008298	AC008298	Drosophil
45	319	18.6	32314	3	AF098993	AF098993	Caenorhab

#### ALIGNMENTS

RESULT 1	AX456998	Sequence 1 from Patent WO228906.	979 bp	DNA	linear	PAT 06-JUL-2002
LOCUS	AX456998	Sequence 1 from Patent WO228906.	979 bp	DNA	linear	PAT 06-JUL-2002
DEFINITION	AX456998	Sequence 1 from Patent WO228906.	979 bp	DNA	linear	PAT 06-JUL-2002
ACCESSION	AX456998	Sequence 1 from Patent WO228906.	979 bp	DNA	linear	PAT 06-JUL-2002
VERSION	AX456998.1	GI:21715789				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	1					
AUTHORS	Kossida, S. and Encinas, J.					
TITLE	Regulation of human sphingosine kinase-like protein					
JOURNAL	Patent: WO 0228906-A 1 11-APR-2002;					
FEATURES	Bayer Aktiengesellschaft (DE)					
source	Location/Qualifiers					
	1..979					
	/organism="Homo sapiens"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:9606"					

#### ORIGIN

Alignment Scores:					
Pred. No.:	2.77e-153	Length:	979		
Score:	1717.00	Matches:	326		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		

```
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-631-958-2 (1-326) x AX456998 (1-979)

Qy 1 ProlyshHisLeuLeuValPheIleAsnProPheGlyGlyGlyGlnGlyLysArgile 20
Db 2 CCAAAGCATTTTACTCGTATTTTATCAACCCGTTTGGAGGAAAAGGACAAGGCAAGCGGATA 61
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40
Db 62 TATGAAAGAAAAGTGGCCACCTGTTACCTTAGCCTCCATCACCCTGACATCATCGGT 121
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 122 AACAAATCTATGTTAACTATGTAGAGTAATTTACTGAACATGCTAATCAGGCCAAGGAG 181
Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
Db 182 ACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGTCGGGGAGATGGT 241
Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 242 ATGTTTCAGCAGGGTCTGCAGCGTCTGATTGGGAGGACGACAGAGAGCGCGGGTTCGAC 301
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db 302 CAGAACCAACCCCGGGCTGTGCTGGTCCCAAGTAGGCTCCGGATTGGAAATCATTTCCGCA 361
Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 362 GGGTCAACGGACTGCTGTGTTACTTCCACCGTGGGACCGACGACGACGAGAACCTCGGCG 421
Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
Db 422 CTGCATATCGTGTGGGAGTTCGCTGGCCATGATGTGTCCTCAGTCCACCAACAGCAGC 481
Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db 482 ACATCTCTTCGCTACTCGCTGCTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAG 541
Qy 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 542 GACAGTGAAGAAAACGGTGGTGGTCTTGCCAGATACGACTTTTTCAGGTTTAAAGACC 601
Qy 201 PheLeuSerHisHisCysTyrGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 602 TTCTCTCTCCACCACTGCTATGAAGGGACAGTGTCCTTCTCCCTGCAACAACACACCGTG 661
Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db 662 GGATCTCCAGGGATAGGAGCCCTGCCGGGACGAGTGTCTTTGTTGACGCAAGCAAGCAAG 721
Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 722 CAGCAGCTGGAGGAGGAGCAGAGAAAGCACTGTATGTTTGGAAAGCTGCGGAGGACGTG 781
Qy 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 782 GAGGAGTGGCAAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAAATGTCTCTGT 841
Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
Db 842 GCTTGTCCCGGAGGCCCGAGGGCTCTCCCGGGCTGCCCCTTGGGAGACGGGTCTCTCT 901
Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
Db 902 GACCTCATCTCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAC 961
Qy 321 ThrAsnGlnAspGln 326
Db 962 ACCAACCGCAGGACCAG 979
```

```
RESULT 2
AX224383 1840 bp DNA linear PAT 10-SEP-2001
LOCUS Sequence 5 from Patent WO0160990.
DEFINITION AX224383
ACCESSION AX224383.1 GI:15554633
VERSION AX224383.1
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Rastelli,L.
TITLE Novel sphingosine kinases
JOURNAL Patent: WO 0160990-A 5 23-AUG-2001;
Curagen Corporation (US); GENENTECH, INC. (US)
FEATURES
source
1. 1840
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: 6.3e-153 Length: 1840
Pred. No.: 1717.00 Matches: 326
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6
US-10-631-958-2 (1-326) x AX224383 (1-1840)
Qy 1 ProlyshHisLeuLeuValPheIleAsnProPheGlyGlyGlyGlnGlyLysArgile 20
Db 265 CCAAAGCATTTTACTCGTATTTTATCAACCCGTTTGGAGGAAAAGGACAAGGCAAGCGGATA 324
Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40
Db 325 TATGAAAGAAAAGTGGCCACCTGTTCCACCTTAGCCTCCATCACCCTGACATCATCGGT 384
Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 385 AACAAATCTATGTTAACTATGTAGAGTAATTTACTGAACATGCTAATCAGGCCAAGGAG 444
Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
Db 445 ACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGGT 504
Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 505 ATGTTTCAGCGAGGTGCTGTCACGGTCTGATTGGGAGGACGACAGAGAGCGCGGGGTTCGAC 564
Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db 565 CAGAACCAACCCCGGGCTGTGCTGGTCCCAAGTACGCTCCAGTATGGAAATCATTTCCGCA 624
Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 625 GGGTCAACGGACTGCGTGTGTTACTCCACCGTGGGACCGACGACGACGAGAACCTCGGCG 684
Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
Db 685 CTGCATATCGTGTGGGACTCGCTGGCCATGGAATGTCCTCAGTCCACCAACAGCAGC 744
Qy 161 ThrLeuLeuArgTyrSerValSerLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db 745 ACATCTCTCGTACTCGGTGTCCTGTCGGGCTACGGCTTCTACGGGAGACATCATCAAG 804
Qy 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 805 GACAGTGAAGAAAACGGTGGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAGACC 864
```



QY 201 PheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220  
 Db 865 TTCTCTCCACACATGCTATGAAGGACAGTGTCTTCTCTCCCTGCACACACACGGTG 924  
 QY 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240  
 Db 925 GGATCTCCAAGGATAGGAAGCCCTGCGCGGACGATGCTTTGTTGTCAGGCAAGCAAG 984  
 QY 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260  
 Db 985 CAGCAGCTCGGAGGAGCAGAGAAAGCACTGTATGTTTGAAGCTGCGGAGGACGTG 1044  
 QY 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaAlaAsnAlaThrAsnMetSerCys 280  
 Db 1045 GAGGAGTGGCAAGTCTCTGTGGAAAGTTCTGGCCATCAATGCCACAAACATGTCCTGT 1104  
 QY 281 AlaCysArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300  
 Db 1105 GCTTGTGCGCGAGGCCAGGCGCTCTCCCGGCTGCCACTTGGGAGCGGCTTCT 1164  
 QY 301 AspLeuLeuLeuLeuLeuLeuLysCysSerArgPheAsnPheLeuArgPheLeuLeuArgHis 320  
 Db 1165 GACCTCATCTCATCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAC 1224  
 QY 321 ThrAsnGlnAspGln 326  
 Db 1225 ACCAACCCAGGACACAC 1242

## RESULT 3

CR456404 1654 bp mRNA linear PRI 25-MAY-2004  
 LOCUS Homo sapiens BK29F11.1 full length open reading frame (ORF) cDNA  
 DEFINITION clone (cDNA) C220RF:pgem.bk29f11.1).  
 ACCESSION CR456404  
 VERSION CR456404.1 GI:47678338  
 KEYWORDS CDNA; chromosome 22; ORF.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1654)  
 AUTHORS Collins, J.E., Wright, C.L., Edwards, C.A., Davis, M.P., Grinham, J.A.,  
 Cole, C.G., Goward, M.E., Aguado, B., Mallya, M., Mokrab, Y.,  
 Huckle, E.J., Beare, D.M. and Dunham, I.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-MAY-2004) Sanger Institute, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: c22g@sanger.ac.uk Manuscript  
 COMMENT Sanger Institute name : pgem.bk29f11.1  
 Homo sapiens cDNA sequence. This sequence was generated as part of  
 The Wellcome Trust Sanger Institute program to isolate cDNA clones  
 representing the full length open reading frame of well annotated  
 protein coding genes on human chromosome 22. For more information  
 see http://www.sanger.ac.uk/HGP/Chr22/.

## FEATURES

Location/Qualifiers  
 1..1654  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="22"  
 /clone="pgem.bk29f11.1"  
 /lab\_host="JM109"  
 1..1654  
 /gene="bk29f11.1"  
 11..1624  
 /gene="bk29f11.1"  
 /codon\_start=1  
 /protein\_id="CAG30290.1"  
 /db\_xref="GI:47678339"  
 /translation="MGATGAAPLQSVKQKCAVSLPARALLRWSPGPGAGA  
 PGADACSVPSVSIIVAEETDVHGKQSGKWKMEKPYFTVHCVKRRRRHWRKWAQV  
 TFWCEPQOLCHLWLOLEMLEKLSKPKHLVFINPFGKGQGRIVYERKVALPFTL  
 ASITTDIIVTEHANAKETLYEINDKVDGIVCVGGDGMFSEVLHGLIGRTQRSAGVD  
 QNHPRAVLVFSSLRIGIIPAGSTDCVSTVGTSDAETSAHLHVVGDLSLAMDVSSVHH

## gene

## CDS

## ORIGIN

Alignment Scores:  
 Pred. No.: 9.96e-146 Length: 1654  
 Score: 1640.50 Matches: 314  
 Percent Similarity: 96.63% Conservative: 1  
 Best Local Similarity: 96.32% Mismatches: 0  
 Query Match: 95.54% Indels: 11  
 Ds: 9 Gaps: 1  
 US-10-631-958-2 (1-326) x CR456404 (1-1654)  
 QY 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20  
 Db 398 CCAAGCAATTTACTGTTATTTATCAACCCGTTGGAGGAAAGGACAGGACAGCGGATA 457  
 QY 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40  
 Db 458 TATGAAAGAAAGTGGCACCACCTGTTCACTTAGCTCCATCACCACCTGACATCATC--- 514  
 QY 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60  
 Db 515 -----GTTACTGAACATGCTAATCAGGCCAAGGAG 544  
 QY 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80  
 Db 545 ACTCTGTATGAGATTAAACATAGACAAATACACGCGCATCGTCTGTGTCGGCGGAGATGCT 604  
 QY 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100  
 Db 605 ATGTTCCAGGAGTGTCTGACGGTCTGATTGGAGGACGACAGAGAGCGCGGGTCCGAC 664  
 QY 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120  
 Db 665 CAGAACCAACCCCGGGCTGTGCTGTCCTCCAGTAGTACCTCCGGATTGGGAATCATTCCTCGCA 724  
 QY 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140  
 Db 725 GGGTCAACGAGTCTGCTGTGTACTTCCACCGTGGGACGACGAGCGAGAACCTTCGGCG 784  
 QY 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160  
 Db 785 CTGCATATCGTTGTTGGGACTCTCGCTGGCCATGATGTGTCTCAGTCCACCAACACAGC 844  
 QY 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180  
 Db 845 ACACCTCTCTCGTACTCTCGTGTCTGCTGCTGGCTACGGCTTCTACGGGACATCATCAAG 904  
 QY 181 AspSerGluLysLysArgTyrLeuGluAlaArgTyrAspPheSerGlyLeuLysThr 200  
 Db 905 GACAGTGAGAAAGAACGGTGTGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACC 964  
 QY 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220  
 Db 965 TTCTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCTCCCTGCACACACACGCTG 1024  
 QY 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240  
 Db 1025 GGATCTCCAAGGATAGGAAGCCCTGCGCGGACGAGTCTTGTGTTGTCAGGCAAGCAAG 1084  
 QY 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260  
 Db 1085 CAGCAGCTGGAGGAGGAGCAGAAAGACACTGTATGTTTGGAGCTGCGGAGGACGTG 1144  
 QY 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280  
 Db 1145 GAGGAGTGGCAAGTCTCTGTGGGAGGTTTCTGGCCATCAATGTCACCAACATGTCCTGT 1204

51  
 43  
 427  
 1314  
 320  
 53

```
Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
    |||||
Db 1205 GCTTGTGCGCGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCT 1264
    |||||
Qy 301 AspleulleLeulleArgLysCysSerArgPheAsnPheLeuArgPheLeuLeuArgHis 320
    |||||
Db 1265 GACCTCATCTCCATCCGGAATGCTCCAGGTTCAAATTTTCTGAGATTTCTCATCAGGCAC 1324
    |||||
Qy 321 ThrAsnGlnGlnAspGln 326
    |||||
Db 1325 ACCAACCCAGGAGACCAG 1342

RESULT 4
LOCUS HSA457828 2042 bp mRNA linear PRI 19-APR-2002
DEFINITION Homo sapiens mRNA for putative lipid kinase (LK4 gene).
ACCESSION AJ457828
VERSION AJ457828.1 GI:20269072
KEYWORDS lipid kinase; LK4 gene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Van Veldhoven,P.P.
TITLE A search for lipid kinases
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2042)
AUTHORS Van Veldhoven,P.P.
TITLE Direct Submission
JOURNAL Submitted (18-APR-2002) Van Veldhoven P.P., Farmakologie,
K.U. Leuven, Herestraat, B-3000 Leuven, BELGIUM

FEATURES
     source          1..2042
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="22"
                     /map="22q13"
     gene            1..2042
                     /gene="LK4"
                     /codon_start=1
                     /name="LK4"
                     /product="putative lipid kinase"
                     /protein_id="CAD29884.1"
                     /db_xref="GI:20269073"
                     /db_xref="GOA:O8TCT0"
                     /translation="MGATGAAPLQSVLWVKQRCVSLPEPARALLRWRSPPGPAGA
                     PGADACSPVSEIITAVETDVHGKGGKWKQKVPYFTVHCVKRARRHWKWAQV
                     TFWCPBEQLCHLWLTLEMLEKLTSPKHLVFINPFGKGQKRIYERKVAFLFTL
                     ASITTDIIVTSHANOAKETLYEINIDKTDGIVCGGDMFSEVLUHLIGRTQRAGVD
                     QNHPRVLVSLRIGTIPAGSTDVCYSTVGTSDAETSAHLIVGDSLAMDVSSVHH
                     NSTLIRYSVLGFGVDIIDKSEKKRWLGARYDFSLGKTFLSHCYEGTVGFLEA
                     QHTVSPDRPCRCAGFCVCRQKQLEOKALYGLEAADVEEQVVCCKFLAIN
                     ATNNSCARRSPRGLSPAHTLGDGSDLLIRKCSRNFNLFRLIRHNTQQDQDPFTFV
                     EYRVKFFQFTSKHMEDESDLKSGGKKRFGHCSSHPSCCTVSNSNWCNDCGEVLHS
                     FAIEVRVHCQLVRLFARGIENRPDSSH"

ORIGIN
Alignment Scores:
Pred. No.:      1,31e-145      Length:      2042
Score:          1640.50      Matches:     314
Percent Similarity: 96.63%      Conservative: 1
Best Local Similarity: 96.32%      Mismatches: 0
Query Match:     95.54%      Indels:      11
DB:              9      Gaps:         1

US-10-631-958-2 (1-326) x HSA457828 (1-2042)

Qy 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyArgIle 20
    |||||
```

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 4171)  
 AUTHORS Ohara, O., Nagase, T. and Nakajima, D.  
 TITLE Novel genes and proteins encoded by the genes  
 JOURNAL Patent: JP 2002345492-A 181 03-DEC-2002;  
 KAZUSA DNA RESEARCH INSTITUTE  
 COMMENT OS Homo sapiens (human)  
 PN JP 2002345492-A/181  
 PD 03-DEC-2002  
 PF 26-FEB-2002 JP 2002049009  
 PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA  
 PC C12N15/09, C07K14/47//A61K31/711.A61K38/00.A61K48/00.A61P25/00,  
 A61P25/14,  
 PC A61P25/18,A61P35/00,C12N15/00,A61K37/02  
 CC Novel genes and proteins encoded by the genes FH Key  
 CDS Location/Qualifiers  
 FT CDS (2)..(1444).  
 FEATURES  
 source  
 1..4171  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3,32e-145 Length: 4171  
 Score: 1640.50 Matches: 314  
 Percent Similarity: 96.63% Conservatives: 1  
 Best Local Similarity: 96.33% Mismatches: 0  
 Query Match: 95.54% Indels: 11  
 DB: 6 Gaps: 1  
 US-10-631-958-2 (1-326) x BD183468 (1-4171)  
 QY 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20  
 DB 221 CCAGAGCATTTACTGCTGTTATATCAACCGTTTGGAGGAAAGGACAAAGGCAAGCGGATA 280  
 QY 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40  
 DB 281 TATGAAGAAAGAGTGGACCACTGTTCACCTTAGCTCCATCACCACATGATCATC--- 337  
 QY 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60  
 DB 338 -----GTTACTGAACATGCTTAATCAGGCCAAGGAG 367  
 QY 61 ThrLeuTyrGluLeuAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80  
 DB 368 ACTCTGTATGAGTTAAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGGT 427  
 QY 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnAtcSerAlaGlyValAsp 100  
 DB 428 ATGTTTCAGCGAGTGTGTCACCGTCTGATTTGGAGAGCGACGAGAGCGCGGGTCCAC 487  
 QY 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120  
 DB 488 CAGAACCCACCCCGGGCTGTGTCGCCAGTAGCCTCCGGATTTGATCATTTCCCGCA 547  
 QY 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140  
 DB 548 GGGTCAACCGGACTGCGGTGTGTACTTCCACCGTGGGCACGACGACGACGAGAACTTCGGCG 607  
 QY 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160  
 DB 608 CTGCATATCGTTGTTGGGACATCGTGGCCATGGATGTGCTCAGTCCACACACACACAGC 667  
 QY 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180  
 DB 668 ACACCTCCTTCGCTACTCCGCTGTCCCTGTGGGCTACGGCTTCTACGGGACATCATCAG 727

QY 181 AspSerGluLysLysArgTTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200  
 DB 728 GACAGTGAAGAAGAACGGTGTGGGTCTTCCAGATACGACTTTTCAGGTTTAAAGACC 787  
 QY 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220  
 DB 788 TTCTCTCTCCACCACTGCTATGAGGACAGTGTCTTCTCTCCCTGCACACACACACGGTG 847  
 QY 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240  
 DB 848 GGATCTCCAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTGAGCAAGCAAG 907  
 QY 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260  
 DB 908 CAGCAGCTGGAGGAGGAGCAGAGAAGACACTGTATGTTTGGAGCTGCGGAGACCGTG 967  
 QY 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280  
 DB 968 GAGGAGTGGCAAGTCGTCTGTGGGAGTTTCTGGCCATCAATGCCACAAACATGTCCTGT 1027  
 QY 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300  
 DB 1028 GCTTGTGCGCGAGCCCGCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCT 1087  
 QY 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320  
 DB 1088 GACCTCATCTCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCAC 1147  
 QY 321 ThrAsnGlnGlnAspGln 326  
 DB 1148 ACCAACCCAGGACCGAC 1165  
 RESULT 6  
 AB051433  
 LOCUS Homo sapiens mRNA for KIAA1646 protein, partial cds. 4171 bp mRNA linear PRI 06-OCT-2001  
 DEFINITION  
 AB051433  
 ACCESSION  
 AB051433.1 GI:13359166  
 VERSION  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Hirose, M., Nagase, T., Murahashi, Y., Kikuno, R. and Ohara, O.  
 TITLE Identification of novel transcribed sequences on human chromosome 22 by expressed sequence tag mapping  
 JOURNAL DNA Res. 8 (1), 1-9 (2001)  
 MEDLINE 21156230  
 PUBMED 11258795  
 REFERENCE 2 (bases 1 to 4171)  
 AUTHORS Ohara, O., Nagase, T. and Kikuno, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (22-NOV-2000) Osamu Ohara, Kazusa DNA Research Institute,  
 Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba  
 292-0812, Japan (E-mail:cdnaifookazusa.or.jp,  
 URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,  
 Fax:81-438-52-3914)  
 FEATURES  
 Location/Qualifiers  
 1..4171  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="hk01650"  
 /note="vector:pBluescriptII SK plus"  
 1..4171  
 /gene="KIAA1646"  
 <1..1447  
 /gene="KIAA1646"  
 /note="Start codon is not identified."  
 /codon\_start=2  
 /product="KIAA1646 protein"  
 /protein\_id="BAB33316.1"

/db\_xref="GI:13359167"  
/translation="IIAVEETDVRHGKQSGKQWMEKPYAFTVHCVKRARRHRWKA  
QVTWCPEQLCHLWLOTREMLKLTSPKHLVFINPFGKGQGGKRIYERKVAPLF  
TLASTITDIIVTEHANQAKETLYEINIDKYDGIVCVGGDGMFSEVLHGLIGRTORSAG  
VDQHPRAVLVPSSLRGIIIPAGSTDVCVSTVGTSDAETSALHIVGDSLAMDVSF  
HNNSTLLRYSVSLIGYFGYDIIKDSKKRWLGARYDFSLKTFLLSHHCYEGTVSFL  
PAQHTVSGPRDKPCRCAGFCVCRQSKQOLEEKALYGLEAAEDVEEWQVCGFLA  
INATNMSCACRRSPRGLSPAHLGDSGLILIRKCSRFNPLREFLIHTNQODQDFT  
FVEYRVKFOFTSKHMEDESDLKEGKKRFGHCSSHPSCCCTVSNSSWNCDEVL  
HSPAIEVRVHCOLVRLFARGIEENPKPDSHS"

## ORIGIN

Alignment Scores:  
Pred. No.: 3.32e-145 Length: 4171  
Score: 1640.50 Matches: 314  
Percent Similarity: 96.63% Conservative: 1  
Best Local Similarity: 96.32% Mismatches: 0  
Query Match: 95.54% Indels: 11  
DB: 9 Gaps: 1

US-10-631-958-2 (1-326) x AB051433 (1-4171)

Qy 1 ProlyshHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20  
Db 221 CCAAAGCATTTACTGGTATTTATCAACCCGTTTGGAGGAAAGGACCAAGCGCGATA 280

Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40  
Db 281 TATGAAAGAAAGTGGACCACTGTTTCACTTAGCTCCATCACCATCATCATC--- 337

Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60  
Db 338 -----GTTACTGAACATGCTAATCATCGGCCCAAGGAG 367

Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80  
Db 368 ACTCTGTATGAGATTAACATAGACAAATACACGCGCATCGTCTGTGTCGGCGGAGATGT 427

Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100  
Db 428 ATGTTTACGAGGAGTCTCAGCGTCTGATTTGGAGGACGAGAGAGCGCGGGGTGCAC 487

Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120  
Db 488 CAGAACCAACCCCGCGCTGTCTGTCGCCAGTAGCTCCGATTTGGAATCATTTCCCGCA 547

Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140  
Db 548 GGGTCAACGGACTGCGTGTGTACTCCACCGTGGGCACACGACGACGAGAACCTCGCGG 607

Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160  
Db 608 CTGCATATCGTTTGGGACCTCGCTGGCCATGGATGTCTCTAGTCCACCAACCAACG 667

Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180  
Db 668 ACACCTCTTCGCTACTCGCTGTCTGTCGGCTACGCGCTTCTACGGGGACATCATCAAG 727

Qy 181 AspSerGluLysLysArgTyrPheLeuAlaArgTyrAspPheSerGlyLeuLysThr 200  
Db 728 GACAGTGAGAGAAACGCTGGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACC 787

Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220  
Db 788 TTCCTCTCCCACTGCTATGAGGGACAGTGCTTCTTCTCCCTGCGACCAACACACGGTG 847

Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240  
Db 848 GGATCTCCAAGGATAGGAAGCCCTCGCGCGGAGGATGCTTTGTTTTCAGGCAAGCAAG 907

Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260  
Db 908 CAGCAGCTGGAGGAGGAGCAGAAAGAACGACTGTATGTTTGGAGAGCTGCGGAGGACGTG 967

Qy 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280  
Db 968 GAGGAGTGGCAAGTCGTCTGTGGAGAGTTTCTGGCCATCAATGCCACAAACATGCTCTGT 1027

Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300  
Db 1028 GCTTGTCCGCGAGCCCGACGAGGCGCTCTCCCGGCTCCACCTTGGGAGACGGGCTTCT 1087

Qy 301 AspleuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320  
Db 1088 GACCTCATCTCTATCCGGAATGCTCCAGGTTCATATTTCTGAGATTTCATCAGGCAC 1147

Qy 321 ThrAsnGlnGlnAspGln 326  
Db 1148 ACCAACCCAGGACGACCAG 1165

RESULT 7  
AX457006 4413 bp DNA linear PAT 06-JUL-2002  
LOCUS Sequence 9 from Patent WO0228906.  
DEFINITION AX457006  
ACCESSION AX457006  
VERSION AX457006.1 GI:21715795  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Kossida, S. and Encinas, J.  
TITLE Regulation of human sphingosine kinase-like protein  
JOURNAL Patent: WO 0228906-A 9 11-APR-2002;  
Bayer Aktiengesellschaft (DE)  
FEATURES  
Location/Qualifiers  
source 1..4413  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
76..1689  
/note="unnamed protein product"  
/codon\_start=1  
/db\_xref="GI:21715796"  
/translation="MGATGAAPLQSLVMVKQORCAVSLPEPARALLRWSPGPGAGA  
PGADACSPVSEIIAIVEETDVHGKQSGKQWMEKPYAFTVHCVKRARRHRWKAQV  
TFWCPEQLCHLWLOTREMLKLTSPKHLVFINPFGKGQGGKRIYERKVAPLFTL  
ASITDIIVTEHANQAKETLYEINIDKYDGIVCVGGDGMFSEVLHGLIGRTORSAGVD  
QNHPRVLVPSSLRGIIIPAGSTDVCVSTVGTSDAETSALHIVGDSLAMDVSFVH  
NSTLLRYSVSLIGYFGYDIIKDSKKRWLGARYDFSLKTFLLSHHCYEGTVSFLA  
QHTVSGPRDKPCRCAGFCVCRQSKQOLEEKALYGLEAAEDVEEWQVCGFLA  
INATNMSCACRRSPRGLSPAHLGDSGLILIRKCSRFNPLREFLIHTNQODQDFTFV  
FVEYRVKFOFTSKHMEDESDLKEGKKRFGHCSSHPSCCCTVSNSSWNCDEVLHS  
PAIEVRVHCOLVRLFARGIEENPKPDSHS"

## ORIGIN

Alignment Scores:  
Pred. No.: 3.58e-145 Length: 4413  
Score: 1640.50 Matches: 314  
Percent Similarity: 96.63% Conservative: 1  
Best Local Similarity: 96.32% Mismatches: 0  
Query Match: 95.54% Indels: 11  
DB: 6 Gaps: 1

US-10-631-958-2 (1-326) x AX457006 (1-4413)

Qy 1 ProlyshHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20  
Db 463 CCAAAGCATTTACTGGTATTTATCAACCCGTTTGGAGGAAAGGACCAAGCGCGATA 522

Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40  
Db 523 TATGAAAGAAAGTGGACCACTGTTTCACTTAGCTCCATCACCATCATCATC--- 579

QY 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60  
Db 580 -----GTTACTGAACATGCTAAATCAGGCCAAGGAG 609  
QY 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80  
Db 610 ACTCTGTATGAGATTAAACATACAGAAATACAGCGCATGCTGTGTGCGCGGAGATGGT 669  
QY 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100  
Db 670 ATGTTTCAAGAGGTGTCGACGGTCTGATTGGAGGACGACAGAGGCGCGGGTTCGAC 729  
QY 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120  
Db 730 CAGAACCCCGGGCTGTGCTGCTGCCAGTAGCTCCGGATTGGAATCATCTCCGCA 789  
QY 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140  
Db 790 GGGTCAACGAGCTGCGTGTGTTACTCCACCGTGGGCACGACGACGAAACCTCGGG 849  
QY 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160  
Db 850 CTCGATATCGTGTGGGACCTCGCTGGCCATGGATGTCTCAGTCCACCAACAGC 909  
QY 161 ThrLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180  
Db 910 ACATCTCTCGTACTCGCTGCTGCTGGCTACGGCTTCTACGGGACATCATCAAG 969  
QY 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200  
Db 970 GACAGTGAAGAAGCGGTGGTGGTCTTGCACATACGACTTTTCAGTTTAAAGACC 1029  
QY 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220  
Db 1030 TTCTCTCTCCACCACTGCTATGAAGGACAGTGTCTCTCTCTCTGACACACACAG 1089  
QY 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240  
Db 1090 GGATCTCCAAGGATAGGAAGCCCTCGCGGCGAGATGCTTTGTTGAGGCAAGCAAG 1149  
QY 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaGluAspVal 260  
Db 1150 CAGCAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1209  
QY 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280  
Db 1210 GAGGAGTGGCAAGTCTGTCTGGGAAGTTCTGGCCATCAATGCCACCAACATGCTGT 1269  
QY 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300  
Db 1270 GCTTGTGCGCGGAGCCCGAGGCGCTCTCCCGGCTGCCACCTTGGGAGACGGGTCTT 1329  
QY 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320  
Db 1330 GACCTCACTCTCATCGGAAATGCTCCAGTTCAATTTTCTGAGATTCTCTCATAGGCAC 1389  
QY 321 ThrAsnGlnGlnAspGln 326  
Db 1390 ACCAACCCAGCAGACCAG 1407  
RESULT 8  
AB079066 4445 bp mRNA linear PRI 27-JUN-2002  
LOCUS Homo sapiens cerk mRNA for ceramide kinase, complete cds.  
DEFINITION  
ACCESSION AB079066  
VERSION  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE  
AUTHORS Sugiyama, M., Kono, K., Liu, H., Shimizu, T., Minekura, H.,

Spiegel, S. and Kohama, T.  
Ceramide Kinase, a Novel Lipid Kinase. MOLECULAR CLONING AND  
FUNCTIONAL CHARACTERIZATION  
J. Biol. Chem. 277 (26), 23294-23300 (2002)  
11956206  
2 (bases 1 to 4445)  
Sugiyama, M., Kono, K., Shimizu, T., Minekura, H., Spiegel, S. and  
Kohama, T.  
Direct Submission  
Submitted (29-JAN-2002) Masako Sugiyama, Sankyo Co., Ltd.,  
Pharmacology and Molecular Biology Research Laboratories; 2-58  
Hiromachi 1-chome, Shinagawa-ku, Tokyo 140-8710, Japan  
(E-mail: meugiyama@shinagawa.sankyo.co.jp, Tel: 81-3-3492-3131,  
Fax: 81-3-5436-8565)  
Location/Qualifiers  
1. .4445  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1. .4445  
/gene="cerk"  
124. .1737  
/gene="cerk"  
/codon\_start=1  
/product="ceramide kinase"  
/protein\_id="BAC01154.1"  
/db\_xref="GI:21624340"  
/translation="MGATGAAPLQSVLMVKQCAVSLEPARALLRWRSPGPGAGA  
PGADACSPVSEIIAVSETDVHGQSGKQMKPKYAFVTHCVKRRRRHMKWAQV  
TFWPEBQLCHLWQTLREMLEKLTSPKHLVFINPEGGKQGRKIVYERKVAFLPTL  
ASITTDIIIVTEHANQAKETLYEINIDKXDVGVCGDGMFSEVLHGLIGRTQRSAGVD  
QNTPRVPLVSLRIGIIPAGSTDCVSTVGTSDAETSAUHI VVGDLSAMDVSSVHH  
NSTLLRYSVLLGYFGYDIIKSEKKRWGLARYDFSLGTLFSLHCHYEGTGSFLPA  
QHTVSGPRDRKPCFAGCVKQSQLEEQKALYGLEAAEDVEEWVCGKFLAIN  
ATNMSCACRSPRGLSPAHLHGDGSDLLILRKSRFNFLRLRHTNQOQFDFTV  
EVYKVKFQFTSKHMEDESDLKESGKKRFGHICSSHPSCCCTVSNSSMNCDEVLHS  
PAIEVRVHCQLVRLPARGIENPKPDSHS"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.61e-145 Length: 4445  
Score: 1640.50 Matches: 314  
Percent Similarity: 96.63% Conservative: 1  
Best Local Similarity: 96.32% Mismatches: 0  
Query Match: 95.54% Indels: 11  
DB: 9 Gaps: 1  
US-10-631-958-2 (1-326) x AB079066 (1-4445)  
QY 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20  
Db 511 CCAAGGCAATTACTGTTATTTATCAACCCGTTTGAGGAGAAAGGACAAAGCGGATA 570  
QY 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIleGly 40  
Db 571 TATGAAAGAAAGTGGCACCACCTTTCACCTTACCTCCATCACCACATCATC--- 627  
QY 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60  
Db 628 -----GTTACTGAACATGCTAAATCAGGCCAAGGAG 657  
QY 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80  
Db 658 ACTCTGTATGAGATTAAACATACAGAAATACAGCGCATGCTGTGTGCGCGGAGATGGT 717  
QY 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100  
Db 718 ATGTTTCAAGAGGTGTCGACGGTCTGATTGGGAGGACGACAGAGCGCGGGTTCGAC 777  
QY 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120  
Db 778 CAGAACCCCGGGCTGTGCTGCTGCCAGTAGCTCCGGATTGGAATCATCTCCGCA 837

```
Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
    |||||
Db 838 GGGTCAACGGACTGGTGTGTACTCCACCGTGGGACACGAGCGAGAAACCTCGGCG 897

Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
    |||||
Db 898 CTGCATATCGTTGTTGGGAGCTCGCTGGCCATGGATGTCTCCTCAGTCCACCAACAGC 957

Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleIys 180
    |||||
Db 958 ACATCTCTTCTGCTACTCCGTGTCCCTGTCTGGGCTACGGCTTCTACGGGACATCATCAAG 1017

Qy 181 AspSerGluIysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
    |||||
Db 1018 GACAGTGAGAGAAACGGTGGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAGACC 1077

Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
    |||||
Db 1078 TTCTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGACACACACACGCTG 1137

Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerIys 240
    |||||
Db 1138 GGATCTCAAGGATAGGAAGCCCTGCCGGGAGGATGCTTTGTGTCAGGCAAGCAAG 1197

Qy 241 GlnGlnLeuGluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
    |||||
Db 1198 CAGCAGCTGGAGGAGGACAGAGAAACACTGTATGTTTGGAAAGCTGCCGGAGGACGTG 1257

Qy 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
    |||||
Db 1258 GAGGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGGCCACAAACATGCTCTGT 1317

Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
    |||||
Db 1318 GCTTGTCCCGAGAGCCCGAGGGGCTCTCCCGGCTGCCCATTTGGGAGACGGGTCTTCT 1377

Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
    |||||
Db 1378 GACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCATCAGGCAC 1437

Qy 321 ThrAsnGlnGlnAspGln 326
    |||||
Db 1438 ACCAACCCAGCAGGACCAG 1455

RESULT 9
LOCUS AR541900 4432 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 148 from patent US 6743619.
ACCESSION AR541900
VERSION AR541900.1 GI:53933980
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4432)
AUTHORS Tang, Y. T., Zhou, P., Goodrich, R., Liu, C., Asundi, V., Ren, F.,
Zhang, J., Zhao, Q. A., Yang, Y., Xue, A. J., Wehrman, T., Wang, J. -R.,
Wang, D., and Drmanac, R. T.
Nucleic acids and polypeptides
JOURNAL Patent: US 6743619-A 148 01-JUN-2004;
FEATURES Location/Qualifiers
source 1. 4432
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1,666-144 Length: 4432
Score: 1633.50 Matches: 313
Percent Similarity: 96.32% Conservative: 1
Best Local Similarity: 96.01% Mismatches: 1
Query Match: 95.14% Indels: 11
DB: 6 Gaps: 1
```

```
US-10-631-958-2 (1-326) x AR541900 (1-4432)
Qy 1 ProlyshisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
    |||||
Db 1591 CCAAAGCATTTTACTGGTATTTTATCAACCCGTTTGGAGAAAAGACAAAGCAAGCGGATA 1650

Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
    |||||
Db 1651 TATGAAAAGAAAGTGGACCACTGTTTCACCTTACCTTACCTTCCATCACCCTGACATCATC- 1707

Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
    |||||
Db 1708 -----GTTACTGAAATGCTAATCAGGCCAAGGAG 1737

Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
    |||||
Db 1738 ACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCTGCTGTGTGCGCGGAGATGGT 1797

Qy 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
    |||||
Db 1798 ATGTTCAAGCGAGGTGCTGCACGGTCTGATTTGGAGGACGACAGAGGACGCGCGGTCGAC 1857

Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
    |||||
Db 1858 CAGAAACACCCCGGGCTGTGCTGCCAGTAGCTCCGGATTGGAAATCATTTCCCGCA 1917

Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
    |||||
Db 1918 GGSTCAACGGACTCGCTGTGTTACTCCACCGTGGCACCGACGACGAGAAACCTCGGCG 1977

Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
    |||||
Db 1978 CTGCATATCGTTGTTGGGACTCTGCTGGCCATGGATGTGCTCTCAGTCCACCAACAGC 2037

Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
    |||||
Db 2038 ACATCTCTCGTACTCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAG 2097

Qy 181 AspSerGluIysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
    |||||
Db 2098 GACAGTCAGAAGAAACGGTGGTGGGCTTGGCCAGATACGACTTTTCAGGTTTAAAGACC 2157

Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
    |||||
Db 2158 TTCTCTCTCCACCACTGCTATGGAGGACAGTGTCTCTTCTCTGACAAACACACCGTG 2217

Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
    |||||
Db 2218 GGATCTCAAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTTGCAGGCAAAAGCAAG 2277

Qy 241 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
    |||||
Db 2278 CAGCAGCTGGAGGAGGACAGAAAGAACACTGTATGTTGGAAAGCTGGCGAGGACGTG 2337

Qy 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
    |||||
Db 2338 GAGGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGGCCATCAATGCCACCAACATGCTCTGT 2397

Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300
    |||||
Db 2398 GCTTGTCCCGGAGGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCT 2457

Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
    |||||
Db 2458 GACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCATCAGGCAC 2517

Qy 321 ThrAsnGlnGlnAspGln 326
    |||||
Db 2518 ACCAACCCAGCAGGACCAG 2535

RESULT 10
BD102675
```



LOCUS BD102675 4463 bp DNA linear PAT 27-AUG-2002  
 DEFINITION Ceramide kinase and DNA thereof.  
 ACCESSION BD102675  
 VERSION BD102675.1 GI:22648249  
 KEYWORDS WO 0196575-A/1.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 SUGIURA, M., KONO, K. and KOHAMA, T.  
 Ceramide kinase and DNA thereof  
 Patent: WO 0196575-A 1 20-DEC-2001;  
 SANKYO CO LTD, MASAKO SUGIURA, KEITA KONO, TAKAFUMI KOHAMA  
 COMMENT OS Homo sapiens (human)  
 PN WO 0196575-A/1  
 PD 20-DEC-2001  
 PF 11-JUN-2001 WO 2001JP004889  
 PR 14-JUN-2000 JP OOP 178039  
 PI MASAKO SUGIURA, KEITA KONO, TAKAFUMI KOHAMA  
 PC C12N15/54, C12N9/12, C12N1/21, C12Q1/48, C07K16/40, A61K31/7125, PC  
 A61K48/00,  
 PC A61P3/04, A61P3/10, A61P7/00, A61P25/00, A61P29/00, A61P31/18, PC  
 A61P35/00  
 CC Ceramide kinase and DNA thereof  
 FH Key Location/Qualifiers  
 FT CDS  
 FEATURES  
 source  
 1. .4463  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,08e-144 Length: 4463  
 Score: 1632.50 Matches: 312  
 Percent Similarity: 96.32% Conservative: 2  
 Best Local Similarity: 95.71% Mismatches: 1  
 Query Match: 95.08% Indels: 11  
 DB: Gaps: 1  
 US-10-631-958-2 (1-326) x BD102675 (1-4463)  
 QY 1 ProLysHisLeuLeuValPheIleIleIleProPheGlyGlyLysGlyGlnGlyLysArgIle 20  
 DB 511 CCAGAGCATTTACTGCTATTTATCAACCGTTTGGAGGAAAGGCAAGGCAAGCGGATA 570  
 QY 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40  
 DB 571 TATGAAGAAGAAAGTGGCCACCTGTTCACCTTAGCCTCCATCACCACCTGACATCATC--- 627  
 QY 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60  
 DB 628 -----GTTACTGAACATGCTTAATCAGGCCAAGGAG 657  
 QY 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80  
 DB 658 ACTCTGTATGAGATTAAACATACACAAATACGACGCATCTGTGTGGCGGAGATGCT 717  
 QY 81 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100  
 DB 718 ATGTTTCAGCGAGGTGTCGACGGTCTGATTGGGAGACGACAGAGGCGCGGGGTCGAC 777  
 QY 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120  
 DB 778 CAGAACCAACCCCGGCTGTGTGTCCTCCAGTAGCCTCCGATTGGAATCATTCCTCCCA 837  
 QY 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140  
 DB 838 GGTCCAACGAGCATGCGTGTGTTACTCCACCGTGGGACCGACGACGAGAACTCGGG 897  
 QY 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerValHisAsnSer 160

DB 898 CTGCATATCGTTGTTGGGACTCGCTGGCCATGAGTGTCTCTAGTCCACACACACAGC 957  
 QY 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180  
 DB 958 ACACCTCTTCTACTCGTCTGCTGCTGGCTACGGCTTCTACGGGGACATCATCAAG 1017  
 QY 181 AspSerGluLysLysArgTyrPheLeuLeuAlaArgTyrAspPheSerGlyLeuLysThr 200  
 DB 1018 GACAGTGAAGAAGAACGGTGGTGTTCGCCAGATACGACTTTTTCAGGTTTAAAGACC 1077  
 QY 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220  
 DB 1078 TTCCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCCCTCCCTGCACACACACGGTG 1137  
 QY 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240  
 DB 1138 GGATCTCCAAGGATAGGAAGCCCTCGCGGCGAGGATGCTTTGTTGTCAGGCAAGCAAG 1197  
 QY 241 GlnGlnLeuGluGluGlnLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260  
 DB 1198 CAGCAGCTGGAGGAGGAGCAGAAAGACACTGTATGGTTTGGAAAGCTGCGGAGGACGTG 1257  
 QY 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280  
 DB 1258 GAGGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGGCCATCAATGCCACCAACATGTCCTGT 1317  
 QY 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 300  
 DB 1318 GCTTGTGCGGAGGAGCCCGGCGCTCTCCCGGCTGCCCACCTGGGAGACGGGTCTTCT 1377  
 QY 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320  
 DB 1378 GACCTCATCTCTATCCGGAATGCTCCAAAGTTCAATTTTCTGAGATTTCTCATCAGGCAC 1437  
 QY 321 ThrAsnGlnGlnAspGln 326  
 DB 1438 ACCAACCCAGCAGGACCAG 1455  
 RESULT 11  
 AB079067  
 LOCUS AB079067 2830 bp mRNA linear ROD 27-JUN-2002  
 DEFINITION Mus musculus cerk mRNA for ceramide kinases, complete cds.  
 ACCESSION AB079067  
 VERSION AB079067.1 GI:21624341  
 KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 SUGIURA, M., KONO, K., LIU, H., SHIMIZUGAWA, T., MINEKURA, H.,  
 SPIEGEL, S. and KOHAMA, T.  
 Ceramide Kinase, a Novel Lipid Kinase. MOLECULAR CLONING AND  
 FUNCTIONAL CHARACTERIZATION  
 J. Biol. Chem. 277 (26), 23294-23300 (2002)  
 REFERENCE 11956206  
 2 (bases 1 to 2830)  
 SUGIURA, M., KONO, K., SHIMIZUGAWA, T., MINEKURA, H., SPIEGEL, S. and  
 KOHAMA, T.  
 Direct Submission  
 Submitted (29-JAN-2002) Masako Sugiura, Sankyo Co., Ltd.,  
 Pharmacology and Molecular Biology Research Laboratories; 2-58  
 Hiromachi 1-chome, Shinagawa-ku, Tokyo 140-8710, Japan  
 (E-mail: meugliura@shina.sankyo.co.jp, Tel: 81-3-3492-3131,  
 Fax: 81-3-5436-8565)  
 FEATURES  
 source  
 location/Qualifiers  
 1. .2830  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 1. .2830  
 /gene="cerk"  
 gene



```

CDS
1..1596
/gene="Cerk"
/codon_start=1
/product="ceramide kinases"
/protein_id="BAC01155.1"
/db_xref="GI:21624342"
/translation="MGAMGAAEPLHSLVWRRRCVLSLEPARALLRWRSPEPGPSA
PGADARSVLVSEITAVEKDEKCEKSHSGRHKMKNPFAFVHRVKRIVRHRWKWARV
TFWGADEQLHLQTLRGLLESLTSRPHLLVFINPFGGKGQKRIYKTVAPLFTL
ASITTEILITHANOAKETLYEINTDSYDGLVCVGDGMSEVLHGLVIGRTQOSAGID
PNHPRVLVPTLRIGIIPAGSTDCVYSTVGTNDAETSALHIIIGSLAIDVSSVHY
HNTLLRSVSLGCFYFGDLIKSEKKRWMLVRYDFSGLTKFLSHQYEGTUSFLPA
QHTLQSRPDNKPAGFCVCRSQKQLEEEKALYGLNAAVEVEVQTCGKFLAN
ATNNKSCGSPGGLSPFAHLKGSDFNLIRKGSRFNLFRLIRHYNQEDQDFTV
EVRVKKFPFTSKVEDEDDNDSKQEKQKFGKI CKDRPSTCTCSASRSSNWCNCGVWHS
PAIEVRVHCQLVRLFARGIEES"

ORIGIN
Alignment Scores:
Pred. No.: 2,01e-132 Length: 2830
Score: 1503.50 Matches: 283
Percent Similarity: 92.64% Conservatives: 19
Best local Similarity: 86.81% Mismatches: 13
Query Match: 87.57% Indels: 11
DB: 10 Gaps: 1

US-10-631-958-2 (1-326) x AB079067 (1-2830)

Qy 1 ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle 20
Db 388 CCGAAGCACTTGTCTGTATTCATCAACCCCTTCGGAGGGAAGGTTCAGGGCAAGCGCATC 447

Qy 21 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleGly 40
Db 448 TATGAAAAACAGCTGGCCCTCTGTTTACCTTGGCTTCCATCCTACCGAGATCATC--- 504

Qy 41 AsnLysPheTyrValAsnTyrValGluValIleThrGluHisAlaAsnGlnAlaLysGlu 60
Db 505 -----ATTACAGAGCATGCCAACCAAGCCAAAGGAG 534

Qy 61 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 80
Db 535 ACTTTATACGAGATCAACACAGACAGCTATGATGGCATCGTGTGGTGGTGGGACGGC 594

Qy 81 MetPheSerGluValIleHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 100
Db 595 ATGTTACGAGAGTCTGCATGGGTGATGGAGACGACGAGCGCTGGTATCGAC 654

Qy 101 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 120
Db 655 CCCATCACCCCGAGCGCTGCTGGTGGCCAGTACCCTCAGATCGGCATCATACCGCA 714

Qy 121 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 140
Db 715 GGGTCCACAGATTGTGTGTGTACTCAACAGTAGGCACAAACGACGACGACATCGGCT 774

Qy 141 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 160
Db 775 TTGCATATCATATTATGGGGACTCATCTGCAATAGACGTGCTCTGTGCATCATCAATAAC 834

Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLys 180
Db 835 ACGTGTCTCGGTACTCGGTTCCTCTGTGGGTACGGTTTCTAGGGGGACTTAATCAAG 894

Qy 181 AspSerGluLysLysArgTrpLeuGlyIleAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 895 GACACGAAAAAGAACCGGTGGATGGCCCTCGCTCGGTATGATTTCTCAGGGTGAAGACC 954

Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 955 TTTCCTCTCATCTACTATGAGAGGACATGTCCTTCTCCACGACACGACACGGTG 1014

Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240

```

```

Db 1015 GGATCTCCACGGGACAATAAACCCCTGCGGGGCTGGGTGCTTCGTGTCAGGACAGCAAG 1074
Qy 241 GlnGlnLeuGluGluGlnGlnLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 1075 CAACAGCTGGAAGAAGAAGAAAGAACCCCTGTATGGCTTGGGAACGCCGGAAGATG 1134
Qy 261 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 1135 GAAGAGTGGCAAGTGACATGTGGGAAGTTCTTGGCCATCAATGCCACCAACATGCTCTGT 1194
Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaHisLeuGlyAspGlySerSer 300
Db 1195 GCTTGTCTCGAGCCCTCGGGGCTGTCCGCCATTTGCCCATCTGGAGATGGGTCTTCT 1254
Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
Db 1255 GACCTCATCTTATCCGGAAGTGTCCAGGTTCAACTTCTTGAGATTTCTCATCCGGAC 1314
Qy 321 ThrAsnGlnGlnAspGln 326
Db 1315 ACGAACCCAGGAGGACCAAG 1332

RESULT 12
AK129416
LOCUS Mus musculus mRNA for mKIAA1646 protein. 3661 bp mRNA linear ROD 21-NOV-2003
ACCESSION AK129416
VERSION AK129416.1 GI:37360495
KEYWORDS F11 CDNA.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Koseki,H., Hiraoka,S.,
Saga,Y., Nagase,T., Ohara,O. and Koga,H.
TITLE Prediction of the coding sequences of mouse homologues of KIAA
gene: III. the complete nucleotide sequences of 500 mouse
KIAA-homologous cDNAs identified by screening of terminal sequences
of cDNA clones randomly sampled from size-fractionated libraries
JOURNAL DNA Res. 10 (4), 167-180 (2003)
MEDLINE 22977043
PUBMED 14621295
REFERENCE 2 (bases 1 to 3661)
AUTHORS Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-2003) Hisashi Koga, Kazusa DNA Research
Institute, Laboratory for Genome Informatics; 2-6-7
Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
(B-mail:mouse@kazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)
COMMENT The CREATE program supported by Japan science and technology
corporation; cDNA full insert sequencing; Kazusa DNA Research
Institute; cDNA library construction, clone selection and 5'- &
3'-end one pass sequencing.
FEATURES
Location/Qualifiers
source
1..3661
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="mp01246"
/tissue_type="embryonic tail"
/notes="vector:modified pBC SK+"
1..3661
/gene="mKIAA1646"
<1..1231
/notes="CDS is predicted by in silico analysis. Start codon
is not identified.
Start codon is not identified."
/codon_start=2
/evidence="not experimental"
/product="mKIAA1646 protein"

```



Db	644	CTGGCCATGGATGTCTCAGTCACCAACACACACACTCTTCGCTACTCGGTGTC	703	
Qy	169	LeuLeuGlyTyrGlyPheTyrGlyAspIleIleIysaspSerGluIyLeYsArGTrpLeu	188	
Db	704	CTGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAAACGGTGGTTT	763	
Qy	189	GlyLeuAlaArgTyrAspPheSerGlyLeuIysThrPheLeuSerHisHisCysTyrGlu	208	
Db	764	GGTCTTGCACATAGCACTTTTCAGGTTTAAGACCTTCCTCTCCACCATGCTATGAA	823	
Qy	209	GlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArglyPro	228	
Db	824	GGGACAGTGTCTCTCTCCCTCCCTGCACACACACGGTGGGATCTCCAAGGATAGGAAGCC	883	
Qy	229	CysArgAlaGlyCysPheValCysArgGlnSerIysGlnGlnLeuGluGluGlnIys	248	
Db	884	TGCCGGGAGGATGCTTTGTTTGCAGGCAAGCAGCAGCTGGAGGAGGACAGAAAG	943	
Qy	249	LysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGluTyrGlnValValCysGly	268	
Db	944	AAAGCAGTGTATGTTTGGAGCTGCGGAGCAGTGGAGAGTGCACAGTCTCTGTGGG	1003	
Qy	269	LysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGly	288	
Db	1004	AAGTTTCTGGCCATCAATGCCACAAACATGTCCTGTGCTGTCTGCGCGGAGCCCGAGGGC	1063	
Qy	289	LeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIleLeuIleArgIysCys	308	
Db	1064	CTCTCCCGGCTGCCCATCTGGGAGACGGGTCTCTTGACCTCATCCTCCGGAATGC	1123	
Qy	309	SerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsnGlnGlnAspGln	326	
Db	1124	TCCAGTTCAATTTCAGATTTCTCATCAGGCACACACACAGCAGGACCAG	1177	
RESULT 14				
BC067255				
LOCUS	BC067255	1772 bp	mrna	linear
DEFINITION	Homo sapiens ceramide kinase, mRNA (cdna clone IMAGE:6185601),			
complete cds.				
ACCESSION	BC067255			
VERSION	BC067255.1	GI:45595582		
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
1 (bases 1 to 1772)				
Strausberg,R.D., Feingold,E.A., Grouse,L.H., Derge,J.G.,				
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,				
Altechul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,				
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J.J., Haieh,P.,				
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,				
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,				
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,				
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,				
Abrahamson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,				
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,				
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,				
Villalón,D.K., Muzny,D.N., Sodergren,E.J., Lu,X., Gibbs,R.A.,				
Fahey,J., Helton,E., Kettner,M., Madan,A., Rodriguez,S.,				
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,				
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,				
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,				
Butterfield,V.S., Krzywinski,M.I., Skalska,U., Smalilus,D.E.,				
Schmerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.				
Generation and initial analysis of more than 15,000 full-length				
human and mouse cDNA sequences				
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
2 (bases 1 to 1772)				
Strausberg,R.				
Direct Submission				
TITLE	Submitted (12-MAR-2004) National Institutes of Health, Mammalian			
JOURNAL				
REFERENCE				
AUTHORS				
Qy	1	ProllyshisLeuLeuValPheIleAsnProPheGlyGlyLys-GlyGlnGlyLysArGIl	20	
Db	447	CCAAAGCATTTACTGGTATTATCAAC-----AAAAGGACAAGGCAAGCGGAT	494	

REMARK COMMENT

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: http://mgc.nci.nih.gov

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. James R. Lupski

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabbu, Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska, Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Series: IRAC Plate: 141 Row: d Column: 4

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21703365

This clone has the following problem: The cds is short compared to the longest cds in the locus.

FEATURES

source

1..1772

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6185601"

/tissue\_type="Peripheral Nervous System, dorsal root ganglion"

/clone\_lib="Lupski\_dorsal\_root\_ganglion"

/lab\_hosts="DH108"

/note="Vector: pCMV-SPORT6"

1..1772

/gene="CERK"

/notes="synonyms: KIAA1646, hCERK, dA59H18.3, DKFZp434E0211, FLJ21430, FLJ23239, dA59H18.2, LK4"

/db\_xref="LocusID:64781"

579..1598

/gene="CERK"

/codon\_start=1

/product="CERK protein"

/protein\_id="AAH67255.1"

/db\_xref="GI:45595583"

/translation="MFSEVLHGLIGRTORSAGVDONHPRAVLVPSSLRIGIIPAGSTD

CVCYVTGTSDAETSAIHIVGDSLANDVSSVFLHNSITLLYSVSLGCGVGDIIKDS

EKKRWGLARYDFSGKLTFLFHSCTYEGTVSFPAQHTVGSFPRKKPCACGFCVRQSK

QOUEEQKALYGLEAAEDVEQVCGFLAINATNMSCACRRSPRLGSPAHLGGG

SSDILLIRKSRFNFLRIHTNQDQDFTFVYVRVKKQFTSKHMEEDSLKE

GKKRFHICSSHPSCCTVSNSWNCDEVLHSPAEIVRVHQCQLVRLFARGIEENPK

PDSHS"

ORIGIN

Alignment Scores:

Pred. No.: 8.4e-128 Length: 1772

Score: 1452.00 Matches: 289

Percent Similarity: 88.38% Conservative: 0

Best Local Similarity: 88.38% Mismatches: 1

Query Match: 84.57% Indels: 38

DB: 9 Gaps: 2

US-10-631-958-2 (1-326) x BC067255 (1-1772)



```
Db 570 CTTTCATATTATTATAGGTGACTGTGACGCTTTAGATGTCCTCATCTGTGTCATCAAAACAAC 629
Qy 161 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyVaspileleLys 180
Db 630 ACGTTTTGAAGTATGCTGTATCATCTGTGGCTATGTTTTTATGGAGATGTTTGAA 689
Qy 181 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 200
Db 690 GACAGTGAAGAAACGGTGGATGGTCCGATGAGATACGACTATTTCAGGCTTCAAGACT 749
Qy 201 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 220
Db 750 TTTCTTTCTCATCATTAATCTATCAAGGAAACAATTTCTTCAACCACGCAAAACACACTG 809
Qy 221 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 240
Db 810 GGATCTCCACGAGATAAGATAGCTGCAGACAGGATGTTATATTTGCAAGGAAGTGAG 869
Qy 241 GlnGlnLeuGluGlnGlnLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 260
Db 870 CAACAACCTGGCAGAACCAACGCAAGAGTGT-----GGATTCAAAACATGAAGAGATGAA 923
Qy 261 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 280
Db 924 GAAGAATCGAAGGTATTATAAGGGAAATTTCTAGCCATCAATGTAGTAAATATATGTGCTGT 983
Qy 281 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyVaspileleLys 300
Db 984 GCCTGTCCACGAGTCCAAAGGTCCTTCACCGCGAGCTCATTTAGCAGATGGTTCAGCT 1043
Qy 301 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 320
Db 1044 GACCTGATTCAGTTTCGGAAATGCTCCAGATTTTGATTTTCTGCGTTATCTTGTGCAGACAT 1103
Qy 321 ThrAsnGlnGlnAspGln 326
Db 1104 ACAACAAGAGATGACCAG 1121
```

Search completed: September 6, 2005, 16:37:43  
Job time : 4353.05 secs

Run on: September 5, 2005, 15:12:23 ; Search time 4448.35 Seconds  
(without alignments)  
10664.096 Million cell updates/

Sequence: 1 accaaagcatttactqatat.....acaccaaccagcagaccacg 979

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

[illegible]

Maximum Match 100%

**E**

1: qb ba:\*

```
2: qb_in: *
3: qb_in: *
```

5: qb ov: \*

```

.7: gb ph: *
32_33

```

ॐ नमो भगवते वासुदेवाय ॥

11: qb stg

13: gb un:

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

score greater than or

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

20

No.	Score	Match Length
1	100	1:00
2	95	1:00
3	90	1:00
4	85	1:00
5	80	1:00
6	75	1:00
7	70	1:00
8	65	1:00
9	60	1:00
10	55	1:00
11	50	1:00
12	45	1:00
13	40	1:00
14	35	1:00
15	30	1:00
16	25	1:00
17	20	1:00
18	15	1:00
19	10	1:00
20	5	1:00

1	979	100.0	97
---	-----	-------	----

3	903	92.2	165
3			165

5	903	92.2	417
---	-----	------	-----

7	903	92.2	441
---	-----	------	-----

9	901.4	92.1	443
---	-------	------	-----

11	826.8	84.5	149
----	-------	------	-----

13	639	65.3	366
----	-----	------	-----

15	464.4	47.4	47
----	-------	------	----

17	362.2	37.0	249
----	-------	------	-----

19	302.2	30.9	32
----	-------	------	----

C 20	203.6	21.4	70850	3	AS043549
C 21	209.8	21.4	181772	2	AC007689
C 22	200.6	20.5	564	6	AC139956
C 23	167.2	17.1	240483	3	AC035955
C 24	167.2	17.1	260331	2	AC135400
C 25	167	17.1	1637	6	AX437003
C 26	164	16.8	219077	2	AC116764
C 27	153	15.6	153	6	AX457004
C 28	152.4	15.6	181772	2	AC007689
C 29	136.8	14.0	522	6	AC124385
C 30	106.2	10.8	325530	3	AC125699
C 31	105.6	10.8	153756	5	BX248240
C 32	91.8	9.4	2333	3	AK112703
C 33	90.4	9.2	1323	5	CR760136
C 34	90.4	9.2	163647	7	AC147916
C 35	90.4	9.2	164496	2	AC147881
C 36	89	9.1	550	6	AX457005
C 37	83	8.5	267546	2	AC125740
C 38	78.6	8.0	1084	6	AR510195
C 39	78.6	8.0	3178	3	BT014925
C 40	77	7.9	1810	3	AY061001
C 41	77	7.9	2064	6	AR509346
C 42	77	7.9	2173	6	CQ608040
C 43	77	7.9	10337	6	AC018030
C 44	77	7.9	22353	2	AC017569
C 45	77	7.9	191009	3	AC007575

## ALIGNMENTS

RESULT 1	
LOCUS	AX456998
DEFINITION	Sequence 1 from Patent WO0228906.
ACCESSION	AX456998
VERSION	AX456998.1
KEYWORDS	GI:21715789
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Crata 1 Mammalia; Eutheria; Primates; Cat Kossida, S. and Encinas, J. Regulation of human sphingosine k Patent: WO 0228906-A 11-APR-200 Bayer Aktiengesellschaft. (DE) Location/Qualifiers 1. .979
FEATURES	/organism="Homo sapiens" /mol_type="unassigned DN /db_xref="taxon:9606"
source	

ORIGIN	Query Match	100.0%;	Score 979;
	Best Local Similarity	100.0%;	Pred. No.
	Matches 979; Conservative	0;	Mismatch
QY	1	ACCAAAGCATTACTGGTATTATTATCAACCGG	
DB	1	ACCAAAGCATTACTGGTATTATTATCAACCGG	
QY	61	ATATGAAAGAAAAGTGGCACCACTGTTTCACCG	
DB	61	ATATGAAAGAAAAGTGGCACCACTGTTTCACCG	
QY	121	TACCAATTCCTATGTTTAACTATGTAGAAAGTA	
DB	121	TACCAATTCCTATGTTTAACTATGTAGAAAGTA	
QY	181	GACTCTGTATGAGATTAAACATAGACAAATAC	
DB	181	GACTCTGTATGAGATTAAACATAGACAAATAC	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	979	100.0	979	6	AX456998	Sequence
2	979	100.0	1840	6	AX224383	Sequence
3	903	92.2	1654	9	CR456404	Homo sapi
4	903	92.2	2042	9	HS457828	Homo sapi
5	903	92.2	4171	6	BD183468	Novel gen
6	903	92.2	4171	9	AB051433	Homo sapi
7	903	92.2	4413	6	AX457006	Sequence
8	903	92.2	4445	9	AB079066	Homo sapi
9	901.4	92.1	4432	6	AB541900	Sequence
10	898.2	91.7	4463	6	BD102675	Ceramide
11	826.8	84.5	1459	6	CQ730476	Sequence
12	764.8	78.1	1772	9	BC067255	Homo sapi
13	639	65.3	3661	10	AK129416	Mus muscu
14	637.4	65.1	2830	10	AB079067	Mus muscu
15	464.4	47.4	474	6	AX457001	Sequence
16	418.6	42.8	1450	5	CR3866590	Gallus ga
17	362.2	37.0	2494	5	BC074350	Xenopus la
18	352.4	36.0	1520	5	BC074110	Xenopus la
19	302.2	30.9	3329	6	AX457002	Sequence







TITLE		Direct Submission	
JOURNAL		Submitted (18-APR-2002) Van Veldhoven P.P., Farmakologie, K.U.Leuven, Herestraat, B-3000 Leuven, BELGIUM	
FEATURES	source	Location/Qualifiers	
		1..2042	
gene	CDS	/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/chromosome="22"	
		/map="22q13"	
		1..2042	
		/gene="LK4"	
		71..1684	
		/gene="LK4"	
		/codon_start=1	
		/product="putative lipid kinase"	
		/protein_id="CAD29884.1"	
		/db_xref="GI:20269073"	
		/db_xref="GOA:Q8TCT0"	
		/db_xref="UniProt/Swiss-Prot:Q8TCT0"	
		/translation="MGATGAEPVLSVLWKQRCVSLRPARALLRWRSPPGAGAPGADACVPVSEIATVRETDVHGKHQSGWKQMEKPYAFTVHCVKRARRHWKQAYTFWCPPEOLCHLWLTQREMLEKUTSRPKHLLVFINPFGKGQGRKRIYERKVAFLPLASITDILIVTHANQAKETLYEINIDKYDGVICVCGDGMFSEVLHGLIGRTQBSAGVDQNHPRVILVPSLRIGITIPAGSTDVCYIVTSGDAETSAIHIVVSGSLAMDVSSVHHNSTLIRYSVLSGLGYFGDIIKSEKKRWGLGARYDFSLKTLFSLHHKVEGTGKFLPANHTVGRSDRKPFCAGFCVRSQKQLEEEOKKALYGLEAAEDVEEVQVVCVKFLAINATNNSCACRSPRGLSPAHLGDGSSDLILIRKCSRNFNLFRLIRHNTQDQDPFTFVEYVRVKFQFTSKMEBEDSLKEGKKRFGHICSSHPSCCCTVSNSSNWCNDEVLHSFAIEVRVHCQLVRLFARGIENPKPDSHS"	
ORIGIN			
Query Match		92.2%; Score 903; DB 9; Length 2042;	
Best Local Similarity		96.6%; Pred. No. 6.3e-229;	
Matches	946; Conservative	0; Mismatches	33; Gaps
1	ACCAAGCATTTACTGGTATTTATCAACCCGTTTGGAGAAAAGGACAAGCAAGCGGAT	60	
457	ACCAAGCATTTACTGGTATTTATCAACCCGTTTGGAGAAAAGGACAAGCAAGCGGAT	516	
61	ATATGAAAGAAAGTGGACCACTGTTCACTTAGCTTCCATCCATCCACTGACATCATCGG	120	
517	ATATGAAAGAAAGTGGACCACTGTTCACTTAGCTTCCATCCATCCACTGACATCATCG-	575	
121	TAACAAAATTCATGTTAACTATGTAGAGTAATTAATCTGAACTGCTAATCAGGCCAAGGA	180	
576	-----TTACTGAACTGCTAATCAGGCCAAGGA	603	
181	GACTCTGTATGAGATTAAACATAGACAAATACGAGGCATCGTCTGTGTCGGCGGAGATGG	240	
604	GACTCTGTATGAGATTAAACATAGACAAATACGAGGCATCGTCTGTGTCGGCGGAGATGG	663	
241	TATGTTTCAGCAGAGTGTGCGAGTCTGATTTGGAGGACCGCAGAGGACGCGCGGGTCTGA	300	
664	TATGTTTCAGCAGAGTGTGCGAGTCTGATTTGGAGGACCGCAGAGGACGCGCGGGTCTGA	723	
301	CCAGAACCAACCCCGGGCTGTGTCGCCAGTAGCCTCCGGATTGGAATCATTTCCCGC	360	
724	CCAGAACCAACCCCGGGCTGTGTCGCCAGTAGCCTCCGGATTGGAATCATTTCCCGC	783	
361	AGGGTCAACGGACTCGTGTGTTACTTCCACCGTGGGCAACGAGGACGAGAAACCTCGGC	420	
784	AGGGTCAACGGACTCGTGTGTTACTTCCACCGTGGGCAACGAGGACGAGAAACCTCGGC	843	
421	GCTGCATATCGTTGTGGGGACTCGTGGCCATGGATGTGTTCTCAGTCCACCAACACAG	480	
844	GCTGCATATCGTTGTGGGGACTCGTGGCCATGGATGTGTTCTCAGTCCACCAACACAG	903	
481	CACACTCTCTCGTACTCCGTGTCCTGCTGGGCTACGGCTTCTACCGGGACATCATCAA	540	
904	CACACTCTCTCGTACTCCGTGTCCTGCTGGGCTACGGCTTCTACCGGGACATCATCAA	963	
541	GGACAGTGAGAGAAACGGTGGTGGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGAC	600	
Db		1023	1083
Qy		601	660
Db		1024	1083
Qy		661	720
Db		1084	1143
Qy		721	780
Db		1144	1203
Qy		781	840
Db		1204	1263
Qy		841	900
Db		1264	1323
Qy		901	960
Db		1324	1383
Qy		961	
Db		1384	
RESULT 5		4171 bp DNA linear	
BD183468	Novel genes and proteins encoded by the genes.		
BD183468	BD183468.1	GI:31875668	
KEYWORDS	JP 2002345492-A/181.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 4171)		
AUTHORS	Ohara, O., Nagase, T. and Nakajima, D.		
TITLE	Novel genes and proteins encoded by the genes		
JOURNAL	Patent: JP 2002345492-A 181 03-DEC-2002;		
COMMENT	KAZUSA DNA RESEARCH INSTITUTE		
	OS	Homo sapiens (human)	
	PN	JP 2002345492-A/181	
	PD	03-DEC-2002	
	PF	26-FEB-2002	JP 2002049009
	PI	OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA	
	PC	C12N15/09; C07K14/47//A61K31/711; A61K38/00; A61K48/00; A61P25/00;	
	PC	A61P25/14;	
	PC	A61P25/18; A61P35/00; C12N15/00; A61K37/02	
	CC	Novel genes and proteins encoded by the genes	PH Key
	Location/Qualifiers		
	FT	CDS	(2)..(1444).
FEATURES	Location/Qualifiers		
source	1..4171		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	92.2%;	Score 903;	DB 6; Length 4171;
Best Local Similarity	96.6%;	Pred. No. 6.9e-229;	
Matches	946; Conservative	0; Mismatches	33; Gaps
1	ACCAAGCATTTACTGGTATTTATCAACCCGTTTGGAGAAAAGGACAAGCAAGCGGAT	60	
220	ACCAAGCATTTACTGGTATTTATCAACCCGTTTGGAGAAAAGGACAAGCAAGCGGAT	279	

```
QY 61 ATATGAAGAAAAGTGGGACCACTGTTTACCTTAGCTCCATCACCACCTGACATCATCGG 120
|||
Db 280 ATATGAAGAAAAGTGGGACCACTGTTTACCTTAGCTCCATCACCACCTGACATCATCG- 338

QY 121 TAAACAAATTCATGTTAACTATGTAGAGTAATTAAGTAACATGCTTAATCAGGCCAAGGA 180
|||
Db 339 -----TTACTGAACATGCTTAATCAGGCCAAGGA 366

QY 181 GACTCTGTATGATTAACATAGACAAATAACGCGGCATCGTCTGTGCGGAGATGG 240
|||
Db 367 GACTCTGTATGATTAACATAGACAAATAACGCGGCATCGTCTGTGCGGAGATGG 426

QY 241 TATGTTTACGAGAGTCTGCACGGTCTGATTGGGAGGACGAGAGGCGCGGGGTGCA 300
|||
Db 427 TATGTTTACGAGAGTCTGCACGGTCTGATTGGGAGGACGAGAGGCGCGGGGTGCA 486

QY 301 CCAGAACCAACCCCGGGCTGCTGGTCCCAAGTAGCTCCGGATTGGAATCATTTCCCGC 360
|||
Db 487 CCAGAACCAACCCCGGGCTGCTGGTCCCAAGTAGCTCCGGATTGGAATCATTTCCCGC 546

QY 361 AGGGTCAACGGACTCGTGTGTTACTTCCACCGTGGGCACCGAGCGAGAAACCTCGGC 420
|||
Db 547 AGGGTCAACGGACTCGTGTGTTACTTCCACCGTGGGCACCGAGCGAGAAACCTCGGC 606

QY 421 GCTGATATCGTGTGTTGGGACTCGTGGCCATGAGTGTCTCAGTCCACCAACAG 480
|||
Db 607 GCTGATATCGTGTGTTGGGACTCGTGGCCATGAGTGTCTCAGTCCACCAACAG 666

QY 481 CACACTCTCTCGCTACTCGTGTCTCTGCTGGGCTACGGCTTCTACGGGACATCATCA 540
|||
Db 667 CACACTCTCTCGCTACTCGTGTCTCTGCTGGGCTACGGCTTCTACGGGACATCATCA 726

QY 541 GGACAGTGCAGAGAAACGGTGTGGGCTTCCAGATACGACTTTTCAGGTTTAAAGAC 600
|||
Db 727 GGACAGTGCAGAGAAACGGTGTGGGCTTCCAGATACGACTTTTCAGGTTTAAAGAC 786

QY 601 CTTCTCTCCACCACTGCTATGAAAGGACAGTGTCTTCTCTCTGCAACACACCGGT 660
|||
Db 787 CTTCTCTCCACCACTGCTATGAAAGGACAGTGTCTTCTCTCTGCAACACACCGGT 846

QY 661 GGGATCTCCAGGGATAGAGACCTCGCGGCGAGGATCTTTGTTCCAGGCAAGCAA 720
|||
Db 847 GGGATCTCCAGGGATAGAGACCTCGCGGCGAGGATCTTTGTTCCAGGCAAGCAA 906

QY 721 GCAGCAGCTGCAGGAGGACAGAGAAAGCACTGTATGTTTGGAGCTGCGGAGACGT 780
|||
Db 907 GCAGCAGCTGCAGGAGGACAGAGAAAGCACTGTATGTTTGGAGCTGCGGAGACGT 966

QY 781 GGAGGAGTGGCAAGTCTGTGTGGGAAAGTTTCTGGCCATCAATGCCAACAATGTCCTG 840
|||
Db 967 GGAGGAGTGGCAAGTCTGTGTGGGAAAGTTTCTGGCCATCAATGCCAACAATGTCCTG 1026

QY 841 TGCTTGTGCGGAGCGCCAGGGGCTCTCCCGGCTGCCACTTGGGAGCGGGTCTTC 900
|||
Db 1027 TGCTTGTGCGGAGCGCCAGGGGCTCTCCCGGCTGCCACTTGGGAGCGGGTCTTC 1086

QY 901 TGACCTCATCTCATCCGGAATGCTCCAGGTTTCAATTTCTGAGATTTCTCATCAGGCA 960
|||
Db 1087 TGACCTCATCTCATCCGGAATGCTCCAGGTTTCAATTTCTGAGATTTCTCATCAGGCA 1146

QY 961 CACCAACCAAGGACCGAG 979
|||
Db 1147 CACCAACCAAGGACCGAG 1165
```

RESULT 6  
AB051433  
LOCUS  
DEFINITION Homo sapiens mRNA for KIAA1646 protein, partial cds.  
ACCESSION AB051433  
VERSION AB051433.1 GI:13359166  
KEYWORDS

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS Hirosewa,M., Nagase,T., Murahashi,Y., Kikuno,R. and Ohara,O.  
TITLE Identification of novel transcribed sequences on human chromosome  
JOURNAL DNA Res. 8 (1), 1-9 (2001)  
MEDLINE 21156230  
PUBMED 11258795  
REFERENCE  
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.  
TITLE Direct Submission  
JOURNAL Submitted (22-NOV-2000) Osamu Ohara, Kazusa DNA Research Institute,  
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba  
292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp.  
URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,  
Fax:81-438-52-3914)

FEATURES  
source

1. .4171  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="hk01650"  
/note="vector:pBluescriptII SK plus"  
1. .4171  
/gene="KIAA1646"  
<1. .1447  
/gene="KIAA1646"  
/note="Start codon is not identified."  
/codon\_start=2  
/product="KIAA1646 protein"  
/protein\_id="BAB33316.1"  
/db\_xref="GI:13359167"  
/translation="IIAIVEDVHGKQSGSKWQMEKPYAFTVHCVRARRHRKWA  
QVTFWEEQLCHLQTLREMLKLTSRPKHLVFPNPGKGQGRKRIYRKKVAPLF  
TLASITTDIIIVTEHANOAKETLYEINDKYDGIIVCGDGMFSEVLHGLIGRTQRSAG  
VDQHPRAVLVPSLRIGIIPAGSTDCVSTGTSDAETSAHLIIVGDSLAMDVSSV  
HNSTLLARYSVSLIGFYGDI IKDSEKRWGLARYDFSLKTLFLSHCYEGTVSFL  
PAQITVGSPPDRKPCFACGFCVRCQKQLEEKALYGLAAEDVEWQVVCCKFLA  
INATNMSCACRRSPRGLSPAHLGDSGLILIRKSRFPFLRLIBHTNQDDQDFLT  
FVEYRVYKQFTSKHMEDESDLDKEGKRKRFHICSHSPSCCTTNSNWCNDEVFL  
HSPAIEVRVHCQVLVRLFARGIEENPKPDSHS"

## gene

## CDS

## ORIGIN

Query Match 92.2%; Score 903; DB 9; Length 4171;  
Best Local Similarity 96.6%; Pred. No. 6.9e-229;  
Matches 946; Conservative 0; Mismatches 0; Indels 33; Gaps 1;  
QY 1 ACCAAAGCATTTACTGGTATTTATCAACCCGTTTGGAGGAAAGGACAAAGCGGAT 60  
Db |||  
QY 220 ACCAAAGCATTTACTGGTATTTATCAACCCGTTTGGAGGAAAGGACAAAGCGGAT 279  
Db |||  
QY 61 ATATGAAGAAAAGTGGCACCACCTGTTCACTTAGCTTCCATCACCACCTGACATCATCGG 120  
Db |||  
QY 280 ATATGAAGAAAAGTGGCACCACCTGTTCACTTAGCTTCCATCACCACCTGACATCATCG- 338  
Db |||  
QY 121 TAACAAAATTCATGTTAACTATGTAGAGTAATTAAGTAACATGCTTAATCAGGCCAAGGA 180  
Db |||  
QY 339 -----TTACTGAACATGCTTAATCAGGCCAAGGA 366  
Db |||  
QY 181 GACTCTGTATGATTAACATAGACAAATAACGCGGCATCGTCTGTGCGGAGATGG 240  
Db |||  
QY 367 GACTCTGTATGATTAACATAGACAAATAACGCGGCATCGTCTGTGCGGAGATGG 426  
Db |||  
QY 241 TATGTTTACGAGAGTCTGCACGGTCTGATTGGGAGGACGAGAGGCGCGGGGTGCA 300  
Db |||  
QY 427 TATGTTTACGAGAGTCTGCACGGTCTGATTGGGAGGACGAGAGGCGCGGGGTGCA 486  
Db |||  
QY 301 CCAGAACCAACCCCGGGCTGCTGGTCCCAAGTAGCTCCGGATTGGAATCATTTCCCGC 360  
Db |||  
QY 487 CCAGAACCAACCCCGGGCTGCTGGTCCCAAGTAGCTCCGGATTGGAATCATTTCCCGC 546  
Db |||

QY	361	AGGGTCAACGAGCTGCGTGTGTTACTCCACCGTGGGACACGAGCGACGAGAAACCTCGGC	420	PGADACSPVSEIIAIVEETDVHGKSGKQWMEKPYAFTVHCVKRARRHRWKAOV
DB	547	AGGGTCAACGAGCTGCGTGTGTTACTCCACCGTGGGACACGAGCGACGAGAAACCTCGGC	606	TFWCPBQLCHLMLOTLREMLEKILSRPKHLVFINPFGKGQGR1YERKVAFLPTL
QY	421	GCTGCATATCGTGTGTTGGGACTCGCTGGCCATGGATGTCTCAGTCCACCAACAG	480	ASITTDILVTEHANQAKELYEINIDKYDGIWCVGDMFSEVLHGLIYGTQRSAGVD
DB	607	GCTGCATATCGTGTGTTGGGACTCGCTGGCCATGGATGTCTCAGTCCACCAACAG	666	QNTPRAVLVFPSSLRIGIIPAGSIDCVICYSTVSGTSDAETSALHIVVGLSLANDVSSVHH
QY	481	CACACTCCTTCGCTACTCCGTTGCTCTCGCTGGGCTACGGCTTCTACGGGACATCATCA	540	NSTLLRYSLLGTYGFGYDIIKDSKKRWGLARDFSLGKTLFLSHCYTGTGVSFLPA
DB	667	CACACTCCTTCGCTACTCCGTTGCTCTCGCTGGGCTACGGCTTCTACGGGACATCATCA	726	QHTVGSFDRDKPCRQFVCRSQQLLEEKKALYGLAAEDFVEHWQVVCVGKFLAIN
QY	541	GGACAGTGAAGAAACCGTGTGGGTCTTTGCCAGATACGACTTTTCAGGTTTAAAGAC	600	ATNNSCACRRSPRGLSPAHLGDGSDLIILIRKCSRFNPLRLIRHTNQDQDFTV
DB	727	GGACAGTGAAGAAACCGTGTGGGTCTTTGCCAGATACGACTTTTCAGGTTTAAAGAC	786	EVTVKXFOFTSKHMEDESDLKEGGKKRGHICSSHPSCCCTVSNSSWNCNDCGBVLHS
QY	601	CTTCTCTCCCAACCACTCTATGAAGGACAGTGTCTTCTCTCCCTGCGACACACCGGT	660	PAIEVRVHCOLVRLFARGIBENPKPDSH*
DB	787	CTTCTCTCTCCCAACCACTCTATGAAGGACAGTGTCTTCTCTCCCTGCGACACACCGGT	846	
QY	661	GGGATCTCCAGGGATAGAGACCTTCCGGGACAGGATGCTTTGTCAGGCAAGCAA	720	
DB	847	GGGATCTCCAGGGATAGAGACCTTCCGGGACAGGATGCTTTGTCAGGCAAGCAA	906	
QY	721	GCAGCAGCTGGAGGAGGACAGAGAAAGCACTGTATGGTTTGAAGCTGGCGAGGACGT	780	
DB	907	GCAGCAGCTGGAGGAGGACAGAGAAAGCACTGTATGGTTTGAAGCTGGCGAGGACGT	966	
QY	781	GGAGAGTGGCAAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCAACAACATGTCTGT	840	
DB	967	GGAGAGTGGCAAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCAACAACATGTCTGT	1026	
QY	841	TGCTTGTCCGCGAGCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGCGGGTCTTC	900	
DB	1027	TGCTTGTCCGCGAGCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGCGGGTCTTC	1086	
QY	901	TGACCTCATCTCATCCGGAATGCTCCAGGTTCATATTTCTTGAGATTCTCATCAGGCA	960	
DB	1087	TGACCTCATCTCATCCGGAATGCTCCAGGTTCATATTTCTTGAGATTCTCATCAGGCA	1146	
QY	961	CACCAACAGCAGGACACAG 979		
DB	1147	CACCAACAGCAGGACACAG 1165		
RESULT 7				
AX457006				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
Homo sapiens (human)				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
1				
Kossida, S. and Encinas, J.				
AUTHORS				
Regulation of human sphingosine kinase-like protein				
TITLE				
Patent: WO 0228906-A 9 11-APR-2002;				
JOURNAL				
Bayer Aktiengesellschaft (DE)				
FEATURES				
Location/Qualifiers				
1..4413				
/organism="Homo sapiens"				
/mol_type="unassigned DNA"				
/db_xref="taxon:9606"				
76..1689				
/note="unnamed protein product"				
/codon_start=1				
/protein_id="CAD38256.1"				
/db_xref="GI:21715796"				
/translation="MGATGAEEPLQSVLWVKQRCVSLPAPALLRWRSPPGAGA				
source				
1..4413				
/organism="Homo sapiens"				
/mol_type="unassigned DNA"				
/db_xref="taxon:9606"				
76..1689				
/note="unnamed protein product"				
/codon_start=1				
/protein_id="CAD38256.1"				
/db_xref="GI:21715796"				
/translation="MGATGAEEPLQSVLWVKQRCVSLPAPALLRWRSPPGAGA				
CDS				
781				
GGAGGAGTGGCAAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCAACAACATGTCTGT				
840				
GGAGGAGTGGCAAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCAACAACATGTCTGT				
1209				
GGAGGAGTGGCAAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCAACAACATGTCTGT				
1268				
GGAGGAGTGGCAAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCAACAACATGTCTGT				
900				
TGCTTGTGGCGGAGCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGCGGGTCTTC				

Db 1269 TGCTTGTCCGCGAGCCCGAGGGCCTCTCCCGGCTGCCACTTGGGAGAGGGGTCTTC 1328  
Qy 901 TGACCTCATCTCTATCCGGAATGCTCAGGTTCAATTTTCGAGATTTCATCAGGCA 960  
Db 1329 TGACCTCATCTCTATCCGGAATGCTCAGGTTCAATTTTCGAGATTTCATCAGGCA 1388  
Qy 961 CACCAACAGCAGGACCG 979  
Db 1389 CACCAACAGCAGGACCG 1407

RESULT 8  
AB079066  
LOCUS AB079066 4445 bp mRNA linear PRI 27-JUN-2002  
DEFINITION Homo sapiens cerk mRNA for ceramide kinase, complete cds.  
ACCESSION AB079066  
VERSION AB079066.1 GI:21624339  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Suglira, M., Kono, K., Liu, H., Shimizugawa, T., Minekura, H.,  
Spiegel, S. and Kohama, T.  
Ceramide kinase, a Novel Lipid Kinase. MOLECULAR CLONING AND  
FUNCTIONAL CHARACTERIZATION  
J. Biol. Chem. 277 (26), 23294-23300 (2002)  
11956206  
2 (bases 1 to 4445)  
Sugira, M., Kono, K., Shimizugawa, T., Minekura, H., Spiegel, S. and  
Kohama, T.  
Direct Submission  
Submitted (29-JAN-2002) Masako Sugira, Sankyo Co., Ltd.,  
Pharmacology and Molecular Biology Research Laboratories; 2-58  
Hiromachi 1-chome, Shinagawa-ku, Tokyo 140-8710, Japan  
(E-mail: msugira@ehina.sankyo.co.jp, Tel: 81-3-3492-3131,  
Fax: 81-3-5436-8565)  
FEATURES  
Location/Qualifiers  
1..4445  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1..4445  
/gene="cerk"  
124..1737  
/gene="cerk"  
/codon\_start=1  
/product="ceramide kinase"  
/protein\_id="BAC01154.1"  
/db\_xref="GI:21624340"  
/translation="MGATGAAPLQSVLWVQRCVSLPAPALLRWRSPPGAGAG  
PGADACSPVSEIIAIVETDVHGKQSGKQKMEKPYATVHCVKRARRHWKWAQV  
TFWCPPEOLCHLQTLREMLEKTSRPKHLVFLNPPGGQGRLYERKVAFLFL  
ASITDIIVTEHANQAKETLYEINIDKYDGLVGVGGDGFSEVLHGLIGRTQASGVD  
QNHPRVLPSSLRIGIIPAGTDCVYSTVGTSDAETSAHLIIVGDSLAMDVSVSH  
NSTLRLYSLLGTFYGDIIKSEKKRWGLARYDFSGKTLFLSHHCYEGTVGFLPA  
QHTVGSPPDRPCRAGCFVCRSQKQLEEQKALYGLEAAEDVEEWQVCGKFLAIN  
ATNMSCACRRSPRGLSPAHLGDSGLILIRKCSRFNRLIRLIRHTNQDQDPFTV  
EYRVKKEQFTSKMEDESDLKSGGKRFHICSSHPSCCTVSNSNWNCDGEVLHS  
PAIEVRVHCQLVRLFARGIENPKPDSHS"

Query Match 92.2%; Score 903; DB 9; Length 4445;  
Best Local Similarity 96.6%; Pred. No. 6.9e-229;  
Matches 946; Conservative 0; Mismatches 0; Indels 33; Gaps 1;  
Qy 1 ACCAAGCATTTACTGGTATTATCAACCCGTTTCGAGAAAGACAAAGCAAGCGGAT 60  
Db 510 ACCAAGCATTTACTGGTATTATCAACCCGTTTCGAGAAAGACAAAGCAAGCGGAT 569  
Qy 61 ATATGAAGAAAGTGGCACCACCTGTTTACCTTAGCTTCCATCACCCTGATCATCGG 120

ORIGIN

Query Match 92.2%; Score 903; DB 9; Length 4445;  
Best Local Similarity 96.6%; Pred. No. 6.9e-229;  
Matches 946; Conservative 0; Mismatches 0; Indels 33; Gaps 1;  
Qy 1 ACCAAGCATTTACTGGTATTATCAACCCGTTTCGAGAAAGACAAAGCAAGCGGAT 60  
Db 510 ACCAAGCATTTACTGGTATTATCAACCCGTTTCGAGAAAGACAAAGCAAGCGGAT 569  
Qy 61 ATATGAAGAAAGTGGCACCACCTGTTTACCTTAGCTTCCATCACCCTGATCATCGG 120

Db 570 ATATGAAGAAAGTGGCACCACCTGTTTACCTTAGCTTCCATCACCCTGATCATCGG - 628  
Qy 121 TAACAAATTTCTATGTTAACTATATGTAAGATTAATTTACTGAACATCTAATCAGGCCAAGGA 180  
Db 629 -----TTACTGAAATGCTAATCAGGCCAAGGA 656  
Qy 181 GACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGG 240  
Db 657 GACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGG 716  
Qy 241 TATGTTTACGAGGTGTCACCGTCTGATTGGGAGGACGCGAGAGGACGCGCGGGGTGGA 300  
Db 717 TATGTTTACGAGGTGTCACCGTCTGATTGGGAGGACGCGAGAGGACGCGCGGGGTGGA 776  
Qy 301 CCAGAACACACCCCGGGCTGTGTCGCCAGTAGCTCCGGAATGGAATCATTTCCCGC 360  
Db 777 CCAGAACACACCCCGGGCTGTGTCGCCAGTAGCTCCGGAATGGAATCATTTCCCGC 836  
Qy 361 AGGCTCAACGACCTGCGTGTGTTACTCCACCGTGGGACGCGAGCGCAAACTCCGC 420  
Db 837 AGGCTCAACGACCTGCGTGTGTTACTCCACCGTGGGACGCGAGCGCAAACTCCGC 896  
Qy 421 GCTCATATCTGTTTGGGGACTCGTGGCCATGGATGTGCTCTCAGTCCACCAACAG 480  
Db 897 GCTCATATCTGTTTGGGGACTCGTGGCCATGGATGTGCTCTCAGTCCACCAACAG 956  
Qy 481 CACACTCTTCGCTACTCCGTCCTGTCGGGCTACGGCTTCTACGGGACATCATCAA 540  
Db 957 CACACTCTTCGCTACTCCGTCCTGTCGGGCTACGGCTTCTACGGGACATCATCAA 1016  
Qy 541 GGACAGTGAAGAAACGGTGGTGGTCTTGCAGATACGACTTTTCAGGTTTAAAGAC 600  
Db 1017 GGACAGTGAAGAAACGGTGGTGGTCTTGCAGATACGACTTTTCAGGTTTAAAGAC 1076  
Qy 601 CTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCTCCCTGCACAAACACACGGT 660  
Db 1077 CTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCTCCCTGCACAAACACACGGT 1136  
Qy 661 GGGATCTCCAAGGATAGGAAGCCCTCGCGGACAGGATGCTTTGTTTCAGCGCAAAAGCAA 720  
Db 1137 GGGATCTCCAAGGATAGGAAGCCCTCGCGGACAGGATGCTTTGTTTCAGCGCAAAAGCAA 1196  
Qy 721 GCACAGCTGGAGGAGGACGAAAGAAAGCACTGTATGTTTGAAGCTCGGAGGACGT 780  
Db 1197 GCACAGCTGGAGGAGGACGAAAGAAAGCACTGTATGTTTGAAGCTCGGAGGACGT 1256  
Qy 781 GGAGAGTGGCAAGTCTGTGTTGGAAAGTTTCTGGCCATCAATGCCACAAACATGTCTCTG 840  
Db 1257 GGAGAGTGGCAAGTCTGTGTTGGAAAGTTTCTGGCCATCAATGCCACAAACATGTCTCTG 1316  
Qy 841 TGCTTGTCCGCGAGCCCCAGGGGCTCTCCCGGCTGCCCATTTGGGAGACGGGTCTTC 900  
Db 1317 TGCTTGTCCGCGAGCCCCAGGGGCTCTCCCGGCTGCCCATTTGGGAGACGGGTCTTC 1376  
Qy 901 TGACCTCATCTCATCGGAAATGCTCAGGTTCAATTTCTGAGATTTCATCAGGCA 960  
Db 1377 TGACCTCATCTCATCGGAAATGCTCAGGTTCAATTTCTGAGATTTCATCAGGCA 1436  
Qy 961 CACCAACAGCAGGACCG 979  
Db 1437 CACCAACAGCAGGACCG 1455

RESULT 9  
AR541900  
LOCUS AR541900 4432 bp DNA linear PAT 08-OCT-2004  
DEFINITION Sequence 148 from patent US 6743619.  
ACCESSION AR541900  
VERSION AR541900.1 GI:53933980  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
Unclassified.





Db 777 CAGAACACCCCGGCTGTCTGGTCCCAAGTAGCCTCCGGATGGAATCATTTCCGCG 836  
Qy 361 AGGCTCAACGGACTCGGTGTGTTACTCCA CCGTGGGCAACGACGACGAGAAACCTCGGC 420  
Db 837 AGGTCCAACGGACTCGGTGTGTTACTCCA CCGTGGGCAACGACGACGAGAAACCTCGGC 896  
Qy 421 GCTGCATATCGTGTGTGGGACTCGCTGGCCATGAGATGTGTCTCAGTCCACCACAAG 480  
Db 897 GCTGCATATCGTGTGTGGGACTCGCTGGCCATGAGATGTGTCTCAGTCCACCACAAG 956  
Qy 481 CACACTCTCTCCCTACTCTCGTGTCCCTGTCTGGGCTACGCTTCTACGGGACATCATCAA 540  
Db 957 CACACTCTCTCCCTACTCTCGTGTCCCTGTCTGGGCTACGCTTCTACGGGACATCATCAA 1016  
Qy 541 GGACAGTGAAGAAACCGTGTGGGTCTTGGCCAGATACGACTTTTTCAGGTTTAAAGAC 600  
Db 1017 GGACAGTGAAGAAACCGTGTGGGTCTTGGCCAGATACGACTTTTTCAGGTTTAAAGAC 1076  
Qy 601 CTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGACACAACACCGGT 660  
Db 1077 CTTCTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGACACAACACCGGT 1136  
Qy 661 GGGATCTCAAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTGACGGCAAGCAA 720  
Db 1137 GGGATCTCAAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTGACGGCAAGCAA 1196  
Qy 721 GCAGCAGCTGGAGGAGGACGAGAAAGCACTGTATGGTTTGGAGCTGCGGAGGACGT 780  
Db 1197 GCAGCAGCTGGAGGAGGACGAGAAAGCACTGTATGGTTTGGAGCTGCGGAGGACGT 1256  
Qy 781 GGAGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCGCAACAATGTCCTG 840  
Db 1257 GGAGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCGCAACAATGTCCTG 1316  
Qy 841 TGCTTGTGCGCGAGCCCGAGGGCTCTCCCGGCTGCCACTTGGGAGAGCGGTCTTC 900  
Db 1317 TGCTTGTGCGCGAGCCCGAGGGCTCTCCCGGCTGCCACTTGGGAGAGCGGTCTTC 1376  
Qy 901 TGACCTCATCTCATCCGGAATGCTCCAGGTTTCAATTTCTGAGATTTCTCATCAGGCA 960  
Db 1377 TGACCTCATCTCATCCGGAATGCTCCAGGTTTCAATTTCTGAGATTTCTCATCAGGCA 1436  
Qy 961 CACCAACGAGGAGGACGAC 979  
Db 1437 CACCAACGAGGAGGACGAC 1455

RESULT 11  
LOCUS CQ730476 1459 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 16410 from Patent WO02068579.  
ACCESSION CQ730476  
VERSION CQ730476.1 GI:42304409  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.  
TITLE Kites, such as nucleic acid arrays, comprising a majority of  
humanexons or transcripts, for detecting expression and other uses  
thereof  
JOURNAL Patent: WO 02068579-A 16410 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES  
source  
1. .1459  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN

Query Match 84.5%; Score 826.8; DB 6; Length 1459;  
Best Local Similarity 99.8%; Pred. No. 1.1e-208;  
Matches 828; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 150 TAATTACTGAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAT 209  
Db 348 TGAATTACTGAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAT 407  
Qy 210 ACAGCGGATCGTCTGTGTGGCCGAGATGTATGTTTCAGCGAGGTGCTGACGGTCTGA 269  
Db 408 ACAGCGGATCGTCTGTGTGGCCGAGATGTATGTTTCAGCGAGGTGCTGACGGTCTGA 467  
Qy 270 TTGGGAGGACGACAGGAGCGCCGGGCTCGACAGAACCCCGGGCTGTGCTGGTCC 329  
Db 468 TTGGGAGGACGACAGGAGCGCCGGGCTCGACAGAACCCCGGGCTGTGCTGGTCC 527  
Qy 330 CCAGTAGCCTCCGATTTGGATCAATCCCGCAGGGTCAACGACCTGCGTGTGTTACTCCA 389  
Db 528 CCAGTAGCCTCCGATTTGGATCAATCCCGCAGGGTCAACGACCTGCGTGTGTTACTCCA 587  
Qy 390 CCGTGGGACACAGCGACGCAAGAAACCTCGGCGCTGCATATCGTTTGGGAGCTCGCTGG 449  
Db 588 CCGTGGGACACAGCGACGCAAGAAACCTCGGCGCTGCATATCGTTTGGGAGCTCGCTGG 647  
Qy 450 CCATGGATGTGTCTCAGTCCACCACAACAGCACTCTTCCGCTACTCTCCGTCCTCC 509  
Db 648 CCATGGATGTGTCTCAGTCCACCACAACAGCACTCTTCCGCTACTCTCCGTCCTCC 707  
Qy 510 TGGCTACCGCTTCTACGGGACATCATCAAGGACAGTGAAGAAACCGTGTGTTGGTGC 569  
Db 708 TGGCTACCGCTTCTACGGGACATCATCAAGGACAGTGAAGAAACCGTGTGTTGGTGC 767  
Qy 570 TTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCCACTCTGCTATGAAGGA 629  
Db 768 TTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCCACTCTGCTATGAAGGA 827  
Qy 630 CAGTGTCTTCTCTCCGACACAACAGCGTGGGATCTCCAGGATAGGAAGCCCTCC 689  
Db 828 CAGTGTCTTCTCTCCGACACAACAGCGTGGGATCTCCAGGATAGGAAGCCCTCC 887  
Qy 690 GGGCAGGATGCTTTGTTTGGAGCAAGCAAGCAGCAGCTGGAGGAGGACAGAAAG 749  
Db 888 GGGCAGGATGCTTTGTTTGGAGCAAGCAAGCAGCAGCTGGAGGAGGACAGAAAG 947  
Qy 750 CACTGTATGTTTTCGAAGCTGCGGAGGACGTTGGAGGAGTGGCAAGTCTGTGGGAAGT 809  
Db 948 CACTGTATGTTTTCGAAGCTGCGGAGGACGTTGGAGGAGTGGCAAGTCTGTGGGAAGT 1007  
Qy 810 TTCTGGCCATCAATGCCACAACATGCTCTGTGTCGCGGAGCCCGAGGGCCCTCT 869  
Db 1008 TTCTGGCCATCAATGCCACAACATGCTCTGTGTCGCGGAGCCCGAGGGCCCTCT 1067  
Qy 870 CCCCCTGCTCCACTTGGGAGACGGGTCTTCTGACCTCATCTCATCGGAAATGCTCCA 929  
Db 1068 CCCCCTGCTCCACTTGGGAGACGGGTCTTCTGACCTCATCTCATCGGAAATGCTCCA 1127  
Qy 930 GGTTCATATTTCTGAGATTTCTCATCGGACACCAACCCAGCAGGACCCAG 979  
Db 1128 GGTTCATATTTCTGAGATTTCTCATCGGACACCAACCCAGCAGGACCCAG 1177  
RESULT 12  
LOCUS BC067255 1772 bp mRNA linear PRI 06-APR-2004  
DEFINITION Homo sapiens ceramide kinase, mRNA (cDNA clone IMAGE:6185601),  
complete cds.  
ACCESSION BC067255  
VERSION BC067255.1 GI:45595582  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.



REFERENCE AUTHORS	1 (bases 1 to 1772) Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heien,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavani,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,D.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Young,A.C., Shvechenko,Y., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Green,E.D., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzyzanski,M.I., Skalska,U., Smalilus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
REFERENCE AUTHORS	2 (bases 1 to 1772) Strausberg,R.
TITLE	Direct Submission
JOURNAL	Submitted (12-MAR-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK COMMENT	NIH-MGC Project URL: <a href="http://mgc.ncbi.nih.gov">http://mgc.ncbi.nih.gov</a> Contact: MGC help desk Email: <a href="mailto:cgapbs-r@mail.nih.gov">cgapbs-r@mail.nih.gov</a> Tissue Procurement: Dr. James R. Lupski CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada <a href="mailto:info@bcgsc.bc.ca">info@bcgsc.bc.ca</a> Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Anara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palquist, Anca Petrescu, Anna Liisa Prabh, Parvaneh Saedi, JR Santos, Angélique Schnerch, Ursula Skalska, Duane Smalilus, Jeff Stott, Miranda Tsai, George Yang, Jacques Schein, Asim Siddiqui, Rob Holt, Marco Marra.
FEATURES source	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: IRAK Plate: 141 Row: d Column: 4 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21703365 This clone has the following problem: The cds is short compared to the longest cds in the locus. Location/Qualifiers 1..1772 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6185601" /tissue_type="Peripheral Nervous System, dorsal root ganglion" /clone_lib="Lupski_dorsal_root_ganglion" /lab_host="DH10B" /note="Vector: PCMV-SPORT6" 1..1772 /gene="CERK" /note="synonyms: KIAA1646, hCERK, dA59H18.3, DKFZp434E0211, FLJ21430, FLJ23239, dA59H18.2, LK4" /db_xref="LocusID:64781" 579..1598 /gene="CERK" /codon_start=1
gene	AK129416 3661 bp mRNA linear ROD 21-NOV-2003
CDS	RESULT 13 AK129416 LOCUS

/product="CERK protein"	
/protein_id="AAH67255.1"	
/db_xref="GI:45595583"	
/db_xref="LocusID:64781"	
/translation="MFSEVLUHGLIGRTQRSAGVDQNHPRAVLIPVSSLRIGLIPAGSTDCVCYVTGTSDAETSAHLIVVGDLSMDVSSVHHNSVLLRYSVSLGYPGDIKIDS EKKRWGLARYDFSGKNTFLSHHCYVCTGFLPAQHTVSGPRDKPCRCAGCFVCRQSK QOLEEOKALYGLEAAEDVEWQVCGKFLAINATNMSCARRSPRLGSPAAHLGQK SSDLLILKCSRNFELRLIHRNQQDQDFTEFVYRVKKFOFTSKHMEDESDLKE GKKRFGHICSSHPSCCTVSNSWNCDEVLHSPALEVVRVHQLVFLFARGIEENPK PDSHS"	
ORIGIN	
Query Match	78.1%; Score 764.8; DB 9; Length 1772;
Best Local Similarity	99.1%; Pred. No. 3.7e-192;
Matches 769; Conservative	0; Mismatches 7; Indels 0; Gaps 0;
Qy	204 ACAATAACGAGCGATCGTCTGTGTCGGCGAGATGTATGTTCAGCGAGGTGTCGACG 263
Db	541 ACTGACATCATCGATCGTCTGTGTCGGCGAGATGTATGTTCAGCGAGGTGTCGACG 600
Qy	264 GTCTGATTGGAGACGACGAGGAGCGCGGGTTCGACCAACACCCCGGCTGTC 323
Db	601 GTCTGATTGGAGACGACGAGGAGCGCGGGTTCGACCAACACCCCGGCTGTC 660
Qy	324 TGGTCCCACTAGTCTCCGGATTGGAATCATTCCTCCGAGGGTCAACGGAATGCTGTGTT 383
Db	661 TGGTCCCACTAGTCTCCGGATTGGAATCATTCCTCCGAGGGTCAACGGAATGCTGTGTT 720
Qy	384 ACTCAACGCTGGGCAACGAGGACGAGAAACCTCGGCGCTGCATATCGTTGTTGGGACT 443
Db	721 ACTCAACGCTGGGCAACGAGGACGAGAAACCTCGGCGCTGCATATCGTTGTTGGGACT 780
Qy	444 CGCTGGCCATGGATGTCTCCTCAGTCCACCAACACACACACTCTTCCTCGCTACTCCGTGT 503
Db	781 CGCTGGCCATGGATGTCTCCTCAGTCCACCAACACACACACTCTTCCTCGCTACTCCGTGT 840
Qy	504 CCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGAGAGAAACCGTGGT 563
Db	841 CCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGAGAGAAACCGTGGT 900
Qy	564 TGGGCTTTGGCAGATACGACTTTTTCAGTTTAAAGACCTTCCTCTCCACCACTGCTATG 623
Db	901 TGGGCTTTGGCAGATACGACTTTTTCAGTTTAAAGACCTTCCTCTCCACCACTGCTATG 960
Qy	624 AAGGACAGTGTCTCTCCTCCGCAACACACAGTGGGATCTCCAAGGATAGGAAGC 683
Db	961 AAGGACAGTGTCTCTCCTCCGCAACACACAGTGGGATCTCCAAGGATAGGAAGC 1020
Qy	684 CCTCGCGGCGAGGATGCTTTGTTTGCAGGCAAGCAAGCAGCAGCTGCGAGGAGGAGCAGA 743
Db	1021 CCTCGCGGCGAGGATGCTTTGTTTGCAGGCAAGCAAGCAGCAGCTGCGAGGAGGAGCAGA 1080
Qy	744 AGAAGACATGTATGTTTGGAGCTCGGAGGACGTGGAGAGTGGCAAGTCTCTGTG 803
Db	1081 AGAAGACATGTATGTTTGGAGCTCGGAGGACGTGGAGAGTGGCAAGTCTCTGTG 1140
Qy	804 GGAAGTTTCTGGCCATCAATGCCAACAATGTCTCTGTCTGTCGCGGAGGCCACAGG 863
Db	1141 GGAAGTTTCTGGCCATCAATGCCAACAATGTCTCTGTCTGTCGCGGAGGCCACAGG 1200
Qy	864 GCCTCTCCCGGCTGCCACTTTGGGAGACGGGTCTTCTGACCTCATCTCATCGGAAAT 923
Db	1201 GCCTCTCCCGGCTGCCACTTTGGGAGACGGGTCTTCTGACCTCATCTCATCGGAAAT 1260
Qy	924 GCTCAGGTTCAATTTTCTGAGATTTCTCATCAGGACACCAACACGAGGACACG 979
Db	1261 GCTCAGGTTCAATTTTCTGAGATTTCTCATCAGGACACCAACACGAGGACACG 1316

```

DEFINITION Mus musculus mRNA for mKIAA1646 protein.
ACCESSION AK129416
VERSION AK129416.1 GI:37360495
KEYWORDS FLI CDNA.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Koseki,H., Hiraoka,S.,
Saga,Y., Nagase,T., Ohara,O. and Koga,H.
TITLE Prediction of the coding sequences of mouse homologues of KIAA
gene: III. the complete nucleotide sequences of 500 mouse
KIAA-homologous cDNAs identified by screening of terminal sequences
of cDNA clones randomly sampled from size-fractionated libraries
JOURNAL DNA Res. 10 (4), 167-180 (2003)
MEDLINE 22977043
PUBMED 14621295
REFERENCE
2 (bases 1 to 3661)
AUTHORS Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-2003) Hiasashi Koga, Kazusa DNA Research
Institute, Laboratory for Genome Informatics; 2-6-7
Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail:mouse@kazusa.or.jp. Tel:81-438-52-3919, Fax:81-438-52-3918)
COMMENT The CREATE program supported by Japan science and technology
corporation; cDNA full insert sequencing; Kazusa DNA Research
Institute; cDNA library construction, clone selection and 5'- &
3'-end one pass sequencing.
FEATURES
source
1..3661
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="mph01246"
/tissue_type="embryonic tail"
/notes="vector:modified pBC SK+"
1..3661
/gene="mKIAA1646"
<1..1231
/gene="mKIAA1646"
/notes="CDS is predicted by in silico analysis. Start codon
is not identified.
Start codon is not identified.
/codon_start=2
/evidence=not experimental
/product="mKIAA1646 protein"
/protein_id="BAC98226.1"
/db_xref="GI:37360496"
/translation="FCFAASRPKHLVFINPFGKGQKRIYKTVAPLFLASITTE
IIITEHANQAKETLYEINTDSYGVICVGGDMFSEVLHVGIRTOQSAGIDPNHRA
VLVPSLIRIGIIPAGSTDCVSVSTGTNDASTALHIIIGDSLADIVSSVHYHNTLLR
YSVSLGFGYDILDKSEKKRWMLVRYDFSLGKTLFUSHOYFETLSFLPAQHTVGS
PRDNKPCPAGCFVCSQKQOLEEKKALYGLNDAEEVEEWQVTCGKFLAINATMSC
ACPRSPGSLFPAHUGDSSDLILIRKSRNFPLRILRHNTQEDQDFTFVEIVRVK
KPHFTSKHVEDNDSDKOEKQKFGKICKDRPSCSTCSASRSSWNCDEGMHSPAIEVR
VHCQLVRLFARGIEES"
ORIGIN
Query Match 65.3%; Score 639; DB 10; Length 3661;
Best Local Similarity 79.8%; Pred. No. 1.2e-158;
Matches 781; Conservative 0; Mismatches 165; Indels 33; Gaps 1;
QY 1 ACCAAGACATTTACTGGTATTATCAACCCGGTTTGGAGGAAAGGACAAAGCGGAT 60
DB 22 ACCGAGACACTTGGTATTATCAACCCCTTCGAGGAAAGGTGAGGCGACGCGAT 81
QY 61 ATATGAAGAAAAGTGGCACCACTGTTCACCTTAGCTCCATCACCCTGATCATCGG 120
DB 82 CTATGAAAAACAGTGGCGCTCTGTTTACCTTGGCTTCCATCACTACGGAGATCATC- 139
QY 121 TAACAAATCTATGTAACTATGTAGAAGTAACTTACTGAACATGTCTATCAGGCCAAGA 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



QY 359 GCAGGTCACGGACTGCGTGTGTTACTCCACCGTGGGCACGACGACGAGAACCTCG 418  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
128 GCAGGTCACAGGACTGCGTGTGTTACTCCACCGTGGGCACGACGACGAGAACCTCG 187  
QY 419 GCGCTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCCTCAGTCCACCAAC 478  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
188 GCGCTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCCTCAGTCCACCAAC 247  
QY 479 AGCACACTCCTTCGCTACTCCGCTGTCCCTGTGGGCTACGGCTTCTACGGGGACATCATC 538  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
248 AGCACACTCCTTCGCTACTCCGCTGTCCCTGTGGGCTACGGCTTCTACGGGGACATCATC 307  
QY 539 AAGGACAGTGAGAAACGGTGGTGGTCTTGCCAGATACGACTTTCAGGTTTAAAG 598  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
308 AAGGACAGTGAGAAACGGTGGTGGTCTTGCCAGATACGACTTTCAGGTTTAAAG 367  
QY 599 ACCTTTCTCTCCACCACTGTCTATGAAGGGACAGTGTCTTCTCCTCCCTGCACACACACG 658  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
368 ACCTTTCTCTCCACCACTGTCTATGAAGGGACAGTGTCTTCTCCTCCCTGCACACACACG 427  
QY 659 GTGGGATCTCCAGGGATAGGAAGCCCTGCCGGGACAGATGCTTTG 704  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
428 GTGGGATCTCCAAGGGATAGGAAGCCCTGCCGGGACAGATGCTTTG 473

Search completed: September 5, 2005, 23:41:36  
Job time : 4461.35 secs

2.

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2005, 15:04:18 ; Search time 564.067 Seconds  
(without alignments)  
10274.357 Million cell updates/sec

Title: US-10-631-958-1  
Perfect score: 979  
Sequence: 1 accaaagattactgtat.....acacaaaccagcaggaccag 979

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	979	100.0	979	6 ABL40822	Abi40822 Human sph
2	979	100.0	1840	5 AAD14426	Aad14426 Human sph
3	903	92.2	1740	8 ADA05679	Ada05679 Human NOV
4	903	92.2	3975	8 ACC70838	Acc70838 Human Sph
5	903	92.2	4231	3 AAA50510	Aaa50510 Human sph
6	903	92.2	4413	6 ABL40828	Abi40828 Human sph
7	903	92.2	4429	12 ADJ96598	Adj96598 Human lip
8	903	92.2	4445	13 ADP55247	Adp55247 Human PRO
9	901.4	92.1	4432	8 ABX70921	Abx70921 Novel hum
10	898.2	91.7	4463	6 ABA96945	Abag96945 Human cer
11	878.4	89.7	1740	12 ADN62844	Adn62844 Human NOV
12	622.2	63.6	2241	5 AAS77728	Aas77728 DNA encod
13	618	63.1	4702	13 ADS10370	Ads10370 Human the
14	598.8	61.2	753	3 AAC76031	Aac76031 Human ORF
15	543	55.5	1570	5 AAS77731	Aas77731 DNA encod
16	543	55.5	2186	5 AAS77730	Aas77730 DNA encod
17	464.4	47.4	474	6 ABL40823	Abi40823 Human sph
18	401.8	41.0	426	5 AAS77727	Aas77727 DNA encod
19	362.8	37.1	411	5 AAS77729	Aas77729 DNA encod
20	302.2	30.9	329	6 ABL40824	Abi40824 Human sph

21	241	24.6	547	3	AAC76592	Aac76592 Human ORF
22	209.8	21.4	15181	4	AAK65588	Aak65588 Human imm
23	209.8	21.4	15185	4	AAK65589	Aak65589 Human imm
24	200.6	20.5	564	8	ABT23453	Abt23453 Immune-re
25	167	17.1	167	6	ABL40825	Abi40825 Human sph
26	153	15.6	153	6	ABL40826	Abi40826 Human sph
27	140	14.3	817	6	ABQ99499	Abq99499 Human cod
28	140	14.3	817	13	ADS11669	Ads11669 Human the
29	136.8	14.0	522	5	AAAD14427	Aad14427 Partial r
30	89	9.1	550	6	ABL40827	Abi40827 Human sph
31	77	7.9	2173	4	ABL25705	Abi25705 Drosophil
32	77	7.9	10337	4	ABL25704	Abi25704 Drosophil
33	69.8	7.1	498	6	ABK33019	Abk33019 DNA encod
34	57.4	5.9	2000	8	ADA71938	Ada71938 Rice gene
35	52.2	5.3	1774	12	ADQ88890	Adq88890 Novel hum
36	51	5.2	763	8	ABT23456	Abt23456 Immune-re
37	49.4	5.0	2000	8	ADA71938	Ada71938 Rice gene
38	43.4	4.4	1050	11	ABD02665	Abd02665 Pseudomon
39	43.4	4.4	1308	11	ABD02575	Abd02575 Pseudomon
40	41.8	4.3	1167	4	AAF81355	Aaf81355 Quorum se
41	41.8	4.3	1167	13	ADS14656	Ads14656 Pseudomon
42	40.4	4.1	1857	6	ABK90199	Abk90199 cDNA enco
43	40.4	4.1	2380	5	AAS14817	Aas14817 Human cDN
44	40.4	4.1	2380	6	ABL59533	Abi59533 Human sph
45	40.4	4.1	2380	10	ADE85298	Ade85298 Farnesyl

ALIGNMENTS

RESULT 1  
ABL40822  
ID ABL40822 standard; cDNA; 979 BP.  
XX ABL40822;  
XX  
DT 03-JUL-2002 (first entry)  
XX  
DE Human sphingosine kinase-like protein encoding cDNA.  
XX  
KW Human sphingosine kinase-like protein; intracellular signalling; gene;  
cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;  
autoimmune disease; rheumatoid arthritis; Parkinson's disease; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key  
CDS Location/Qualifiers  
FT 2..789  
FT /\*tag= a  
FT /product= "sphingosine kinase-like protein"  
FT /note= "start and stop codons are not indicated"  
XX  
PN WO200228906-A2.  
XX  
PD 11-APR-2002.  
XX  
PF 05-OCT-2001; 2001WO-EP011516.  
XX  
PR 06-OCT-2000; 2000US-02380005P.  
PR 23-AUG-2001; 2001US-0314113P.  
XX  
XX (FARB ) BAYER AG.  
XX  
XX Kossida S, Encinas J;  
XX  
XX WPI; 2002-340094/37.  
XX  
XX P-PSDB; ABB07854.  
XX  
XX New reagent for modulating the activity of sphingosine kinase-like  
protein polypeptide or polynucleotide and treating cancer, asthma,  
allergy, an autoimmune disease, or a central or peripheral nervous system  
disorder.

PS	Claim 1; Fig 1; 120pp; English.	
XX	The invention relates to a human sphingosine kinase-like protein. The	
CC	polypeptide can be expressed by standard recombinant methodology. The	
CC	sphingosine kinase-like protein and gene can be used to regulate and	
CC	intracellular signalling and consequently cell proliferation and	
CC	apoptosis. Such regulation is useful for treating cancer, allergies (e.g.	
CC	asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and	
CC	peripheral nervous system disorders (e.g. Parkinson's disease). The	
CC	present sequence represents the human sphingosine kinase-like protein	
CC	encoding cDNA	
XX		
SQ	Sequence 979 BP; 237 A; 259 C; 271 G; 212 T; 0 U; 0 Other;	
	Query Match 100.0%; Score 979; DB 6; Length 979;	
	Best Local Similarity 100.0%; Pred. No. 2.3e-281;	
	Matches 979; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ACCAAGCATTTACTGTTATTTATCAACCCGTTTGGAGGAAAGGACAAGGCAAGCGGAT 60	
DB	1 ACCAAGCATTTACTGTTATTTATCAACCCGTTTGGAGGAAAGGACNAGGCAAGCGGAT 60	
QY	61 ATATGAAAGAAAGTGGCACCACCTGTTCCACCTTAGCCTCCATCCACTGACATATCGG 120	
DB	61 ATATGAAAGAAAGTGGCACCACCTGTTCCACCTTAGCCTCCATCCACTGACATATCGG 120	
QY	121 TAACAAATTCATGTTACTAGTGTAGAGTAATTAAGTGAATGCTTAATCAGGCGCAAGGA 180	
DB	121 TAACAAATTCATGTTAACTATGTAGAAAGTAATTAAGTGAATGCTTAATCAGGCGCAAGGA 180	
QY	181 GACTCTGTATGATTAACATAGACAAATACGAGGCATCGTCTGTGCGCGGAGATGG 240	
DB	181 GACTCTGTATGATTAACATAGACAAATACGAGGCATCGTCTGTGCGCGGAGATGG 240	
QY	241 TATGTTACGAGGTGCTGCACGGTCTGATTGGGAGGACGAGAGGCGCGGGTCTGA 300	
DB	241 TATGTTACGAGGTGCTGCACGGTCTGATTGGGAGGACGAGAGGCGCGGGTCTGA 300	
QY	301 CCAGAACACCCCGGGTGTGCTGGTCCCGAGTAGCTCCGGATTGGGAATCATTCGCGC 360	
DB	301 CCAGAACACCCCGGGTGTGCTGGTCCCGAGTAGCTCCGGATTGGGAATCATTCGCGC 360	
QY	361 AGGGTCAACGGAAGTGGTGTGTTACTCCACCGTGGGACACGAGCGAGAAACCTCGGC 420	
DB	361 AGGGTCAACGGAAGTGGTGTGTTACTCCACCGTGGGACACGAGCGAGAAACCTCGGC 420	
QY	421 GCTCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCAGTCCACCAACAG 480	
DB	421 GCTCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCAGTCCACCAACAG 480	
QY	481 CACACTCTTCGCTACTCCGTGTCCTGCTGGGTACCGCTTCTACGGGACATCATCAA 540	
DB	481 CACACTCTTCGCTACTCCGTGTCCTGCTGGGTACCGCTTCTACGGGACATCATCAA 540	
QY	541 GGACAGTGCAGAGAAACGGTGGTGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGAC 600	
DB	541 GGACAGTGCAGAGAAACGGTGGTGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGAC 600	
QY	601 TTCTCTTCCACCACTGCTATGAAAGGACAGTGTCTTCTCTCTGTCACAAACACAGGT 660	
DB	601 TTCTCTTCCACCACTGCTATGAAAGGACAGTGTCTTCTCTCTGTCACAAACACAGGT 660	
QY	661 GGGATCTCCAGGGATAGAGCCCTCGCGGACAGGATGCTTTGTTTCAGCGCAAGCAA 720	
DB	661 GGGATCTCCAGGGATAGAGCCCTCGCGGACAGGATGCTTTGTTTCAGCGCAAGCAA 720	
QY	721 GCAGCAGTGGAGGAGGACGAAAGCACTGTATGTTTGGAGCTGGGAGGACGT 780	
DB	721 GCAGCAGTGGAGGAGGACGAAAGCACTGTATGTTTGGAGCTGGGAGGACGT 780	
QY	781 GGAGAGTGGCAAGTCTGTTGGGAAGTTTCTGGCCATCAATGCCAACAATGTCCTG 840	
DB	781 GGAGAGTGGCAAGTCTGTTGGGAAGTTTCTGGCCATCAATGCCAACAATGTCCTG 840	

QY	841 TGCTTGTGCGCGGAGCCCGAGGGCCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTC 900	
DB	841 TGCTTGTGCGCGGAGCCCGAGGGCCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTC 900	
QY	901 TGACCTCATCTCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 960	
DB	901 TGACCTCATCTCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 960	
QY	961 CACCAACCAGCAGGACCAG 979	
DB	961 CACCAACCAGCAGGACCAG 979	
RESULT 2		
AAD14426		
ID	AAD14426 standard; cDNA; 1840 BP.	
XX	AC AAD14426;	
XX	AC AC	
DT	01-NOV-2001 (first entry)	
XX	Human sphingosine kinase (SphK) cDNA #2.	
DE	Human; sphingosine kinase; SphK; restenosis; ischaemia; gene therapy;	
XX	antisense therapy; cancer; sphingolipid; signalling molecule; apoptosis;	
KW	cytostatic; tumour necrosis factor-alpha; TNF; atherosclerosis; lymphoma;	
KW	leukaemia; vasotropic; cell proliferative disorder; vascular disease; ss.	
XX	Homo sapiens.	
OS		
XX	Key Location/Qualifiers	
FH	109..1524	
FT	/*tag= a	
FT	/product= "Human sphingosine kinase (SphK) protein #2"	
FT		
XX	WO200160990-A2.	
XX	23-AUG-2001.	
XX	14-FEB-2001; 2001WO-US004789.	
PF	14-FEB-2000; 2000US-0182360P.	
PR	22-MAR-2000; 2000US-0191261P.	
XX	(CURA-) CURAGEN CORP.	
PA	(GETH ) GENENTECH INC.	
XX	Rastelli L;	
XX	WPI; 2001-514770/56.	
DR	P-PSDB; AAE07884.	
XX	An isolated Sphingosine kinase polypeptide useful for treating a SphK-	
FT	associated disorder especially cancer, restenosis or ischemia in a human.	
XX	Claim 8; Page 94-95; 107pp; English.	
XX	The present invention relates to sphingosine kinase (SphK) polypeptides	
CC	and nucleic acids encoding them. SphK is useful for treating a SphK-	
CC	associated disorder especially cancers such as leukaemia, lymphoma,	
CC	ovarian, breast, lung, colon, testicular, stomach and skin,	
CC	atherosclerosis, restenosis or ischaemia and cell proliferative disease	
CC	or disorder associated with vascular diseases. SphK gene is used in gene	
CC	therapy and antisense-therapy. Sphingolipids serving as signalling	
CC	molecules, have recently emerged as regulators of cell growth,	
CC	differentiation, diverse cell phenotypes and cell death. Activation of	
CC	SphK by tumour necrosis factor (TNF)-alpha inhibits apoptosis in human	
CC	endothelial cells. The present sequence is human sphingosine kinase	
CC	(SphK) cDNA	
XX	Sequence 1840 BP; 459 A; 468 C; 503 G; 410 T; 0 U; 0 Other;	
SQ		





PT New NOVX polypeptides and nucleic acids, useful for diagnosing,  
PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,  
PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or  
XX pharmacogenomics.

PS Claim 20; Page 134; 586pp; English.

XX  
CC The present invention describes NOVX proteins, where X can be 1 to 55  
CC (e.g. NOVX). Also described: (1) a composition comprising a polypeptide  
CC described above and a carrier; (2) a kit comprising, in one or more  
CC containers, the composition described above; (3) an isolated nucleic acid  
CC molecule which encodes a NOVX protein of the invention; (4) a vector  
CC comprising the nucleic acid molecule described above; (5) a cell  
CC comprising the above vector; (6) an antibody that immunospecifically  
CC binds to the polypeptide described above; (7) methods for determining the  
CC presence or amount of the above polypeptide or nucleic acid molecule in a  
CC sample; (8) methods for determining the presence of or predisposition to  
CC a disease associated with altered levels of expression of the above  
CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a  
CC method of identifying an agent that binds to the polypeptide described  
CC above; (10) a method for identifying a potential therapeutic agent for  
CC use in treating a pathology that is related to an aberrant expression or  
CC aberrant physiological interactions of the polypeptide; (11) a method of  
CC screening for a modulator of activity or of latency or predisposition to  
CC a pathology associated with the polypeptide; (12) a method for modulating  
CC the activity of the polypeptide described above; (13) methods of treating  
CC or preventing a pathology associated with the above polypeptide in a  
CC mammal; and (14) a method for producing the above polypeptide. NOVX  
CC sequences have antidiabetic, anorectic, antibacterial, virucide,  
CC immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian  
CC and antilipemic activities, and can be used in gene therapy. The  
CC polypeptide is useful in manufacturing a medicament for treating a  
CC acid molecule may be used to diagnose, treat or prevent metabolic  
CC disorders such as diabetes or obesity, infections, cachexia, cancer,  
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's  
CC disease, immune disorders, haematopoietic disorders and various  
CC dyslipidaemias. The nucleic acids can also be used as hybridisation  
CC probes, in chromosome mapping, tissue typing, preventive medicine and  
CC pharmacogenomics. The present sequence encodes a human NOVX protein from  
XX the present invention.

SQ Sequence 1740 BP; 389 A; 476 C; 530 G; 345 T; 0 U; 0 Other;

Query Match 92.2%; Score 903; DB 8; Length 1740;  
Best Local Similarity 96.6%; Pred. No. 1.5e-258;  
Matches 946; Conservative 0; Mismatches 0; Indels 33; Gaps 1;  
QY 1 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGAAAGGACAAAGCGGAT 60  
DB 462 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGAAAGGACAAAGCGGAT 521  
QY 61 ATATGAAAGAAAGTGGCACCCTGTTCACTTACCTTCCATCAACCTGACATCATCGG 120  
DB 522 ATATGAAAGAAAGTGGCACCCTGTTCACTTACCTTCCATCAACCTGACATCATCG- 580  
QY 121 TAACAAATTCATTGTTACTATGTAGAAGTAATTACTGAACATCTAATCAGGCCAAGGA 180  
DB 581 -----TACTGAACATGCTAATCAGGCCAAGGA 608  
QY 181 GACTCTGTATGAGATTAAACATAGACAAATACGAGCATCGTGTGTCGGCGGATGG 240  
DB 609 GACTCTGTATGAGATTAAACATAGACAAATACGAGCATCGTGTGTCGGCGGATGG 668  
QY 241 TATGTTACGAGGTGCTGCACCGTCTCATTTGGAGGACGACAGAGCGCCGGGTGGA 300  
DB 669 TATGTTACGAGGTGCTGCACCGTCTCATTTGGAGGACGACAGAGCGCCGGGTGGA 728  
QY 301 CCAGAACACCCCGGGTGTGTCCTCCAGTAGCTCCGGATTGGAATCATTTCCCGC 360  
DB 729 CCAGAACACCCCGGGTGTGTCCTCCAGTAGCTCCGGATTGGAATCATTTCCCGC 788  
QY 361 AGGGTCAACGGACTGCGTGTGTTACTCCACCGTGGGCCACGACGCGAGAAACCTCGGC 420

DB 789 AGGGTCAACGGACTGCGTGTGTTACTCCACCGTGGGCACGACGCGAGAAACCTCGGC 848  
QY 421 GCTGCATATCGTTGTTGGGACATCGCTGGGCATGGATGTCTCATGCCACAACAG 480  
DB 849 GCTGCATATCGTTGTTGGGACTCGCTGGGCATGGATGTCTCATGCCACAACAG 908  
QY 481 CACACTCTTCCTGCTACTCCGCTGCTCCGCTGGGCTACGGCTCTACGGGACATCATCA 540  
DB 909 CACACTCTTCCTGCTACTCCGCTGCTCCGCTGGGCTACGGCTCTACGGGACATCATCA 968  
QY 541 GGACAGTGAAGAAACCGTGGTGGTCTTTGGCAGATACGACTTTTCAGGTTTAAAGAC 600  
DB 969 GGACAGTGAAGAAACCGTGGTGGTCTTTGGCAGATACGACTTTTCAGGTTTAAAGAC 1028  
QY 601 CTTCTCTCCACCATCTGCTATGAAGGACAGTGTCTCTCTCCCTGCACACACACGGT 660  
DB 1029 CTTCTCTCCACCATCTGCTATGAAGGACAGTGTCTCTCTCCCTGCACACACACGGT 1088  
QY 661 GGGATCTCCAAGGATAGGAAGCCCTCCGCGGACAGTGTCTTTGTCAGGCAAGCA 720  
DB 1089 GGGATCTCCAAGGATAGGAAGCCCTCCGCGGACAGTGTCTTTGTCAGGCAAGCA 1148  
QY 721 GCACAGCTGGAGGAGCAGAAAGCACTGTATGTGTTTGGAACTCGCGAGGACGT 780  
DB 1149 GCACAGCTGGAGGAGCAGAAAGCACTGTATGTGTTTGGAACTCGCGAGGACGT 1208  
QY 781 G3AGGAGTGGCAAGTCTCTGTGGGAAATTTCTGGCCATCAATGCCACAACATGTCTG 840  
DB 1209 G3AGGAGTGGCAAGTCTCTGTGGGAAATTTCTGGCCATCAATGCCACAACATGTCTG 1268  
QY 841 TGCTTGTGCGCGGAGCCCAAGGGCCCTCTCCCGGCTGCCACTTTGGGAGACGGGTCTTC 900  
DB 1269 TGCTTGTGCGCGGAGCCCAAGGGCCCTCTCCCGGCTGCCACTTTGGGAGACGGGTCTTC 1328  
QY 901 TGACCTCATCTCTCATCGGAAATGCTCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 960  
DB 1329 TGACCTCATCTCTCATCGGAAATGCTCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 1388  
QY 961 CACCAACGAGGACGACAG 979  
DB 1389 CACCAACGAGGACGACAG 1407  
RESULT 4  
ACC70838  
ID ACC70838 standard; DNA; 3975 BP.  
XX AC ACC70838;  
XX AC  
DT 20-NOV-2003 (first entry)  
XX DE Human Sphingosine kinase 4 coding sequence.  
XX DE Human; enzyme; haemostatic; sphingosine kinase 4; SPHK4;  
XX KW platelet transfusion; platelet stabiliser; gene; ds.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
XX CDS 1..1251  
XX /\*tag= a  
XX /product= "Sphingosine kinase 4"  
XX WO2003031627-A1.  
XX PN 17-APR-2003.  
XX PD 28-SEP-2001; 2001WO-JP008537.  
XX PF 28-SEP-2001; 2001WO-JP008537.  
XX PR (HOKK-) HOKKAIDO TECHNOLOGY LICENSING OFFICE CO.  
XX PA



CC individuals. Host cells expressing SK can be used in drug screening.  
CC Human SK specific antibodies, inhibitors, ligands or their analogues are  
CC useful as bioactive agents to treat inflammation or disease including  
CC viral, bacterial or fungal infections, allergic responses, mechanical  
CC injury associated with trauma, hereditary diseases, lymphoma or  
CC carcinoma, and other conditions with activate the genes of kidney, lung,  
CC heart, lymphoid or tissues of the nervous system  
XX  
SQ Sequence 4231 BP; 1022 A; 1021 C; 1086 G; 1067 T; 0 U; 35 Other;

Query Match 92.2%; Score 903; DB 3; Length 4231;  
Best Local Similarity 96.6%; Pred. No. 2.4e-258;  
Matches 946; Conservative 0; Mismatches 0; Indels 33; Gaps 1;

Qy 1 ACCAAGCATTTACTGGTATTATACACCCGTTTGGAGGAAGGACAGCGGAT 60  
Db |||||  
Qy 226 ACCAAGCATTTACTGGTATTATACACCCGTTTGGAGGAAGGACAGCGGAT 285  
Db |||||  
Qy 61 ATATGAAGAAAGTGCGCACCACTGTTTACCTTAGCCCTCCATCACCCTGACATCATCGG 120  
Db |||||  
Qy 286 ATATGAAGAAAGTGCGCACCACTGTTTACCTTAGCCCTCCATCACCCTGACATCATCG- 344  
Db |||||  
Qy 121 TAACAAATTCATGTTTAACTAGTAGAAGTAATTAATCTGAACATGCTAATCAGGCCAAGGA 180  
Db |||||  
Qy 345 -----TTACTGAACATGCTAATCAGGCCAAGGA 372  
Db |||||  
Qy 181 GACTCTGTATCAGATTACATAGACAATACGCGGCATCTCTGTGTCGCGGAGATGG 240  
Db |||||  
Qy 373 GACTCTGTATCAGATTACATAGACAATACGCGGCATCTCTGTGTCGCGGAGATGG 432  
Db |||||  
Qy 241 TATGTTACGAGGTGCTGCAAGTCTGATTTGGAGGACGCGAGGCGCGGGTCTGA 300  
Db |||||  
Qy 433 TATGTTACGAGGTGCTGCAAGTCTGATTTGGAGGACGCGAGGCGCGGGTCTGA 492  
Db |||||  
Qy 301 CCAGAACCAACCCCGGGTGTGTCGTCCTCCAGTAGCCTCCGATGGAATCATTCCTCGC 360  
Db |||||  
Qy 493 CCAGAACCAACCCCGGGTGTGTCGTCCTCCAGTAGCCTCCGATGGAATCATTCCTCGC 552  
Db |||||  
Qy 361 AGGGTCAACGACGCTGCTGTTACTCCACCGTGGGACGCGAGCGGAGAACTCGGC 420  
Db |||||  
Qy 553 AGGGTCAACGACGCTGCTGTTACTCCACCGTGGGACGCGAGCGGAGAACTCGGC 612  
Db |||||  
Qy 421 GCTCATATCGTGTGTTGGGACTCGCTGGCCATGGAATGTCCTCAGTCCACCAACAG 480  
Db |||||  
Qy 613 GCTCATATCGTGTGTTGGGACTCGCTGGCCATGGAATGTCCTCAGTCCACCAACAG 672  
Db |||||  
Qy 481 CACACTCTTCGCTACTCTCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATCAAA 540  
Db |||||  
Qy 673 CACACTCTTCGCTACTCTCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATCAAA 732  
Db |||||  
Qy 541 GCACAGTCAGAGAAACGGTGTGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGAC 600  
Db |||||  
Qy 733 GCACAGTCAGAGAAACGGTGTGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGAC 792  
Db |||||  
Qy 601 CTTCTCTCCCAACCACTGCTATGAAGGACAGTGTCTTCTCCTCCCTGCACAAACAGGT 660  
Db |||||  
Qy 793 CTTCTCTCCCAACCACTGCTATGAAGGACAGTGTCTTCTCCTCCCTGCACAAACAGGT 852  
Db |||||  
Qy 661 GGGATCTCCAAAGGATAGAAAGCCCTGCGGCGAGGATGCTTTGTCAGCGGAAAGCAA 720  
Db |||||  
Qy 853 GGGATCTCCAAAGGATAGAAAGCCCTGCGGCGAGGATGCTTTGTCAGCGGAAAGCAA 912  
Db |||||  
Qy 721 GCAGCAGCTGAGGAGGAGGAGAGAAAGACTGTATGTTTGAAGCTGCGGAGGACGT 780  
Db |||||  
Qy 913 GCAGCAGCTGAGGAGGAGGAGAGAAAGACTGTATGTTTGAAGCTGCGGAGGACGT 972  
Db |||||  
Qy 781 GGAGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGGCCATCAATGCACAAACATGTCTTG 840  
Db |||||  
Qy 973 GGAGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGGCCATCAATGCACAAACATGTCTTG 1032  
Db |||||  
Qy 841 TGCTTGTGCGCGAGCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTC 900  
Db |||||  
Qy 1033 TGCTTGTGCGCGAGCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTC 1092  
Db |||||

Qy 901 TGACCTCATCTCATCGGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGCA 960  
Db |||||  
Qy 1093 TGACCTCATCTCATCGGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGCA 1152  
Db |||||  
Qy 961 CACCAACCAGCAGGACCAG 979  
Db |||||  
Qy 1153 CACCAACCAGCAGGACCAG 1171  
Db |||||  
RESULT 6  
ID ABL40828 standard; cDNA; 4413 BP.  
XX ABL40828;  
AC ABL40828;  
XX AC  
XX DT  
XX 03-JUL-2002 (first entry)  
DE Human sphingosine kinase-like protein encoding cDNA.  
XX Human sphingosine kinase-like protein; intracellular signalling; gene;  
KW cell proliferation; apoptosis; cancer; allergy; cytostatic; asthma;  
KW autoimmune disease; rheumatoid arthritis; Parkinson's disease; ss.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FT CDS 1..1689  
FT /\*tag= a  
FT /product= "sphingosine kinase-like protein"  
FT /note= "see ABB07857"  
FT 76..1689  
FT /\*tag= b  
FT /product= "sphingosine kinase-like protein"  
FT /note= "see ABB07856"  
XX W0200228906-A2.  
XX 11-APR-2002.  
XX 05-OCT-2001; 2001WO-EP011516.  
XX 06-OCT-2000; 2000US-0238005P.  
XX 23-AUG-2001; 2001US-0314113P.  
XX (FARB ) BAYER AG.  
XX Kossida S, Encinas J;  
XX WPI; 2002-340094/37.  
XX P-PSDB; ABB07856, ABB07857.  
XX New reagent for modulating the activity of sphingosine kinase-like  
XX protein polypeptide or polynucleotide and treating cancer, asthma,  
XX allergy, an autoimmune disease, or a central or peripheral nervous system  
XX disorder.  
XX Claim 1; Fig 9; 120pp; English.  
XX The invention relates to a human sphingosine kinase-like protein. The  
XX polypeptide can be expressed by standard recombinant methodology. The  
XX sphingosine kinase-like protein and gene can be used to regulate  
XX intracellular signalling and consequently cell proliferation and  
XX apoptosis. Such regulation is useful for treating cancer, allergies (e.g.  
XX asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and  
XX peripheral nervous system disorders (e.g. Parkinson's disease). The  
XX present sequence represents the human sphingosine kinase-like protein  
XX encoding cDNA  
XX Sequence 4413 BP; 1020 A; 1107 C; 1207 G; 1079 T; 0 U; 0 Other;  
Query Match 92.2%; Score 903; DB 6; Length 4413;  
Best Local Similarity 96.6%; Pred. No. 2.5e-258;

	Matches	946;	Conservative	0;	Mismatches	0;	Indels	33;	Gaps	1;
Qy	1	ACCAAAGCATTTTACTGTGTTATTTATCAACCCGCTTTGGAGGAAAGGACAAGCGGAT	60							
Db	462	ACCAAAGCATTTTACTGTGTTATTTATCAACCCGCTTTGGAGGAAAGGACAAGCGGAT	521							
Qy	61	ATATGAAGAAAGTGGCACACCTGTTCACCTTAGCCTCCATCACCACTGACATCATCGG	120							
Db	522	ATA7GAAGAAAGTGGCACACCTGTTCACCTTAGCCTCCATCACCACTGACATCATCG-	580							
Qy	121	TAACAAAATTTCTATGTTAACTATGTAGAAGTAAATTTACTGAAACATGCTAATCAGGCGCAAGGA	180							
Db	581	-----TTACTGAAATGCTAATCAGGCCAAGGA	608							
Qy	181	GACTCTGTATGAGATTAACTAGACAAATACGACGGCATCGTCTGTGTCTGGCGGAGATGG	240							
Db	609	GACTCTGTATGAGATTAACTAGACAAATACGACGGCATCGTCTGTGTCTGGCGGAGATGG	668							
Qy	241	TATGTTCAAGCGAGTGTGCAACGCTGTGATTGGGAGGACGACAGGAGAGCCCGGGGTGCA	300							
Db	669	TATGTTCAAGCGAGTGTGCAACGCTGTGATTGGGAGGACGACAGGAGAGCCCGGGGTGCA	728							
Qy	301	CCAGAACCAACCCCGGGCTGTGCTGGTCCCGAGTAGCTCCGGATTGGAATCATTCGCCG	360							
Db	729	CCAGAACCAACCCCGGGCTGTGCTGGTCCCGAGTAGCTCCGGATTGGAATCATTCGCCG	788							
Qy	361	AGGCTCAACGGACTGCGTGTGTTACTCCAAGTGGGACACAGCGACGCAAACTCTGGC	420							
Db	789	AGGCTCAACGGACTGCGTGTGTTACTCCAAGTGGGACACAGCGACGCAAACTCTGGC	848							
Qy	421	GCTGCATATCGTTCTGGGGACTCGCTGGCCATGGATGTGCTCAGTCACACCAACACAG	480							
Db	849	GCTGCATATCGTTCTGGGGACTCGCTGGCCATGGATGTGCTCAGTCACACCAACACAG	908							
Qy	481	CACACTCTTTGGTACTCCGTTGCCCTGTGCTGGGTACGGCTTCTACGGGACATCATCAA	540							
Db	909	CACACTCTTTGGTACTCCGTTGCCCTGTGCTGGGTACGGCTTCTACGGGACATCATCAA	968							
Qy	541	GGACAGTGAGAGAAACGGTGTGGGCTTTGCCAGATACGATTTTCAGGTTTAAAGAC	600							
Db	969	GGACAGTGAGAGAAACGGTGTGGGCTTTGCCAGATACGATTTTCAGGTTTAAAGAC	1028							
Qy	601	CTTCCTCTCCACCACTGCTATGAAGGACAGTGTCTTCCTCCCTGCACAAACACACGGT	660							
Db	1029	CTTCCTCTCCACCACTGCTATGAAGGACAGTGTCTTCCTCCCTGCACAAACACACGGT	1088							
Qy	661	GGGATCTCCAAAGGATAGGAAGCCCTGCGGGGACAGATGCTTTGTTTTCAGGCAAAAGCAA	720							
Db	1089	GGGATCTCCAAAGGATAGGAAGCCCTGCGGGGACAGATGCTTTGTTTTCAGGCAAAAGCAA	1148							
Qy	721	GCACAGCTGGAGGAGGACAGAGAAGCACTGATGTTTTCGGAAGCTCGGGAGGACGT	780							
Db	1149	GCACAGCTGGAGGAGGACAGAGAAGCACTGATGTTTTCGGAAGCTCGGGAGGACGT	1208							
Qy	781	GGAGGAGTGGCAAGTCTGTGGGAAAGTTTCTGGCCATCAATGCCACAAACATGTCCTG	840							
Db	1209	GGAGGAGTGGCAAGTCTGTGGGAAAGTTTCTGGCCATCAATGCCACAAACATGTCCTG	1268							
Qy	841	TGCTTGTGCGCGGAGCCGACGGGCGCTCTCCCGGCTGCCCATCTTGGGAGACGGGTCTTC	900							
Db	1269	TGCTTGTGCGCGGAGCCGACGGGCGCTCTCCCGGCTGCCCATCTTGGGAGACGGGTCTTC	1328							
Qy	901	TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA	960							
Db	1329	TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA	1388							
Qy	961	CACCAACGACGAGACCAG	979							
Db	1389	CACCAACGACGAGACCAG	1407							

ID	ADJ96598	standard; DNA; 4429 BP.
AC	ADJ96598;	
XX		
DT	06-MAY-2004	(first entry)
XX		
DE	Human lipid kinase KIAA1646	DNA SeqID 55.
XX		
KW	gene; ds; kinase; human; SNP; single nucleotide polymorphism;	
KW	tyrosine protein kinase; serine/threonine protein kinase; PKK; STK;	
KW	gene therapy; cancer; immune-related disease; cardiovascular disease;	
KW	brain; neuronal associated disease; metabolic; inflammatory disorder;	
KW	cytostatic; neuroprotective; immunomodulator; antiinflammatory;	
XX	lipid kinase; KIAA1646.	
XX		
OS	Homo sapiens.	
OS	39.	
XX		
PH	Key	Location/Qualifiers
FT	variation	replace(2391,g)
FT		/*tag= a
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	replace(2577,g)
FT		/*tag= b
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	replace(3020,c)
FT		/*tag= c
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	replace(3769,c)
FT		/*tag= d
FT		/standard_name= "Single nucleotide polymorphism"
FT	variation	replace(4272,g)
FT		/*tag= e
FT		/standard_name= "Single nucleotide polymorphism"
XX		
PN	WO2004006838-A2.	
XX		
PD	22-JAN-2004.	
XX		
PF	15-JUL-2003; 2003WO-US021730.	
XX		
PR	15-JUL-2002; 2002US-0395632P.	
XX		
PA	(SUGE-) SUGEN INC.	
XX		
PI	Whyte D, Manning G, Caenepeel S;	
XX		
DR	WPI; 2004-122753/12.	
DR	P-PSDB; ADJ96664.	
XX		
PT	New nucleic acid molecule encoding a kinase polypeptide, useful for	
PT	preparing a composition for treating diseases or disorders, e.g., cancer,	
PT	or neurological, immunological or inflammatory disorders.	
XX		
PS	Example 1; SEQ ID NO 55; 366pp; English.	
XX		
CC	This invention relates to a novel isolated, enriched or purified nucleic	
CC	acid molecule that encodes a kinase polypeptide. Specifically, it relates	
CC	to human tyrosine and serine/threonine protein kinases (PK's and STK's),	
CC	as well as protein kinase-like enzymes. The present invention describes	
CC	screening methods to identify agonists, antagonists and antibodies that	
CC	can be used to modulate the activity or function of the mammalian kinase	
CC	enzymes. As such, these compositions can be used for gene therapy	
CC	purposes to treat diseases or disorders including cancer, immune-related	
CC	diseases, cardiovascular disease, brain or neuronal associated disease,	
CC	metabolic and inflammatory disorders. Accordingly, they exhibit	
CC	cytostatic, neuroprotective, immunomodulator and antiinflammatory	
CC	activities. This polynucleotide sequence is a human kinase DNA sequence	
CC	of the invention.	
XX		
SQ	Sequence 4429 BP; 1020 A; 1115 C; 1213 G; 1081 T; 0 U; 0 Other;	





Best Local Similarity 96.6%; Pred. No. 2.5e-258;		Matches 946; Conservative 0; Mismatches 0; Indels 33; Gaps 1;	
QY	1	ACCAAGACATTTACTGGTATTTATCAACCCGTTTGGAGAAAGGACAAAGCAAGCGGAT	60
DB	510	ACCAAGACATTTACTGGTATTTATCAACCCGTTTGGAGAAAGGACAAAGCAAGCGGAT	569
QY	61	ATATGAAGAAAGTGGACCACTGTGTACCTTAGCTCCATCACCCTGATCATCGG	120
DB	570	ATATGAAGAAAGTGGACCACTGTGTACCTTAGCTCCATCACCCTGATCATCGG	628
QY	121	TAACAAATTCATGTATTAAGTATTAAGTAACTGAACTGTAATCTAGGCGCAAGGA	180
DB	629	-----TTACTGAACATGCTAATCAGGCCAAGGA	656
QY	181	GACTCTGTATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGCGGAGATGG	240
DB	657	GACTCTGTATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGCGGAGATGG	716
QY	241	TATGTTACGAGGTGCTGCACGCTGTGATTGGGAGGACGACAGAGAGCGCGGGTGA	300
DB	717	TATGTTACGAGGTGCTGCACGCTGTGATTGGGAGGACGACAGAGAGCGCGGGTGA	776
QY	301	CCAGAACACACCCCGGGTGTCTGCTCCAGTAGCCTCCGGATTGGAATCATTTCCGC	360
DB	777	CCAGAACACACCCCGGGTGTCTGCTCCAGTAGCCTCCGGATTGGAATCATTTCCGC	836
QY	361	AGGGTCAACGACTCGTGTGTACTCCACCGTGGGACACGACGAGCAAACTCCGC	420
DB	837	AGGGTCAACGACTCGTGTGTACTCCACCGTGGGACACGACGAGCAAACTCCGC	896
QY	421	GCTCATATCGTTGTGGGACTCGCTGGCCATGATGTCTCAGTCCACCAACAG	480
DB	897	GCTCATATCGTTGTGGGACTCGCTGGCCATGATGTCTCAGTCCACCAACAG	956
QY	481	CACACTCTCTCGTACTCGTGTCTGCTGGGCTACGGCTTCTACGGGACATCAA	540
DB	957	CACACTCTCTCGTACTCGTGTCTGCTGGGCTACGGCTTCTACGGGACATCAA	1016
QY	541	GCACAGTGAAGAAACGGTGTGGGTCTGCCAGATACGACTTTTCAGGTTTAAAGAC	600
DB	1017	GCACAGTGAAGAAACGGTGTGGGTCTGCCAGATACGACTTTTCAGGTTTAAAGAC	1076
QY	601	TTCTCTCTCCCACTGCTATGAAGGACAGTGTCTTCTCTCCGACACACACGGT	660
DB	1077	TTCTCTCTCCCACTGCTATGAAGGACAGTGTCTTCTCTCCGACACACACGGT	1136
QY	661	GGGATCTCCAGGGATAGAACCTCTCGGGCAGGATGCTTTGTTGCGAGCAAGCAA	720
DB	1137	GGGATCTCCAGGGATAGAACCTCTCGGGCAGGATGCTTTGTTGCGAGCAAGCAA	1196
QY	721	GCAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGGAAAGCTGCGGAGCGT	780
DB	1197	GCAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGGAAAGCTGCGGAGCGT	1256
QY	781	GGAGGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCCAACAATCTCTG	840
DB	1257	GGAGGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCCAACAATCTCTG	1316
QY	841	TGCTTGTGCGGAGGCCCGGGGCTCTCCCGGCTGCCACTTGGGAGAGCGGTCTTC	900
DB	1317	TGCTTGTGCGGAGGCCCGGGGCTCTCCCGGCTGCCACTTGGGAGAGCGGTCTTC	1376
QY	901	TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTCTCATCAGCA	960
DB	1377	TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTCTCATCAGCA	1436
QY	961	CACCAACCAAGGAGGACAG 979	
DB	1437	CACCAACCAAGGAGGACAG 1455	

ABX70921	ABX70921 standard; cDNA; 4432 BP.
ID	AC
XX	ABX70921;
XX	AC
DT	05-MAR-2003 (first entry)
XX	XX
DE	Novel human cDNA sequence #146.
XX	XX
KW	Human; gene; ss; nervous system disorder; peripheral neuropathy;
KW	Huntington's disease; amyotrophic lateral sclerosis; haemophilia;
KW	neurodegenerative disease; Parkinson's disease; Alzheimer's disease;
KW	autoimmune disease; systemic lupus erythematosus; rheumatoid arthritis;
KW	insulin-dependent diabetes mellitus; anaemia; thrombocytopaenia; wound;
KW	ulcer; burn; bone disorder; osteoporosis; osteoarthritis; stroke;
KW	fibrosis; reperfusion injury; infection; allergic rhinitis; asthma;
KW	coagulation disorder; cancer; tumour; inflammatory disease; septic shock;
KW	Crohn's disease; anaphylaxis; proliferation; chemotactic; chemokinetic;
KW	differentiation; stem cell growth factor; haematopoiesis; haemostatic; antiinflammatory; expressed sequence tag; EST.
XX	XX
OS	Homo sapiens.
XX	XX
PN	W0200281731-A2.
XX	XX
PD	17-OCT-2002.
XX	XX
PF	29-JAN-2002; 2002WO-US001222.
XX	XX
PR	30-JAN-2001; 2001US-00774528.
XX	XX
PA	(HYSE-) HYSEQ INC.
XX	(GOOD/) GOODRICH R W.
PI	Tang TY, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
XX	Xue AJ, Yang Y, Wehrman T, Wang J, Wang D, Dmanac RT;
DR	WPI; 2003-058563/05.
XX	XX
PT	Novel polypeptide useful for treating neurodegenerative diseases, myeloid
PT	or lymphoid cell disorders, bone disorders, mechanical and traumatic
PT	disorders, coagulation disorders, and inflammatory diseases.
XX	XX
PS	Claim 1; Page; 612pp; English.
XX	XX
CC	This invention relates to the cDNA sequences encoding an isolated novel
CC	human polypeptide. The protein encoded by the nucleic acid of the
CC	invention is useful for treating central and peripheral nervous system
CC	diseases (e.g. peripheral neuropathy, Huntington's disease, amyotrophic
CC	lateral sclerosis); neurodegenerative diseases (e.g. Parkinson's disease,
CC	Alzheimer's disease); autoimmune disease (e.g. systemic lupus
CC	erythematosus, rheumatoid arthritis, insulin-dependent diabetes mellitus)
CC	; myeloid or lymphoid cell disorders (e.g. anaemia and thrombocytopaenia)
CC	; wounds, ulcers, burns; bone disorders (e.g. osteoporosis,
CC	osteoarthritis); mechanical and traumatic disorders (e.g. stroke, head
CC	trauma); lung or liver fibrosis; reperfusion injury in various tissues;
CC	bacterial, viral or fungal infections; allergic conditions such as
CC	allergic rhinitis, asthma; coagulation disorders (e.g. haemophilia);
CC	cancer and tumours; and inflammatory diseases (e.g. septic shock, Crohn's
CC	disease, anaphylaxis). The protein may be used to inhibit the growth,
CC	infection or function of infectious agents such as bacteria, fungi,
CC	viruses, or to effect bodily characteristics, biorhythms or circadian
CC	cycles of rhythms. The protein may also have
CC	proliferation/differentiation, stem cell growth factor, haematopoiesis
CC	regulation, immune stimulating or suppressing, chemotactic/chemokinetic,
CC	haemostatic and thrombolytic, receptor/ligand, and antiinflammatory
CC	activities. The cDNA sequences of the invention are useful for expressing
CC	recombinant protein for analysis. The present sequence represents a novel
CC	human cDNA sequence of the invention, this sequence is an expressed
CC	sequence tag (EST) and was identified using subtractive hybridisation
XX	XX
SQ	Sequence 4432 BP; 970 A; 1122 C; 1282 G; 1058 T; 0 U; 0 Other;



```
Query Match 92.1%; Score 901.4; DB 8; Length 4432;
Best Local Similarity 96.5%; Pred. No. 7.5e-258;
Matches 945; Conservative 0; Mismatches 1; Indels 33; Gaps 1;

QY 1 ACCAAGCATTTACTGTTATTTATCAACCCGTTTGGAGGAAAAGACAAAGCAAGCGGAT 60
DB 1590 ACCAAGCATTTACTGTTATTTATCAACCCGTTTGGAGGAAAAGACAAAGCAAGCGGAT 1649

QY 61 ATATGAAAGAAAAGTGGCACCACCTGTTTCACTTGTAGCTTCCATCACCCTGACATCATCGG 120
DB 1650 ATATGAAAGAAAAGTGGCACCACCTGTTTCACTTGTAGCTTCCATCACCCTGACATCATCG- 1708

QY 121 TAACAAATTCATGTTAACTATGTAGAAAGTAATTACTGAACATGCTAATCAGGCCAAGGA 180
DB 1709 -----TTACTGAACATGCTAATCAGGCCAAGGA 1736

QY 181 GACTCTGTATGAGATTAAATAGACAAATATAGACGGCATCGTCTGTGTCGGCGAGATGG 240
DB 1737 GACTCTGTATGAGATTAAATAGACAAATATAGACGGCATCGTCTGTGTCGGCGAGATGG 1796

QY 241 TATGTTTACGCGAGGTGCTGCGGTCTGATTTGGGAGGACGAGAGGCGCGCGGGTCCA 300
DB 1797 TATGTTTACGCGAGGTGCTGCGGTCTGATTTGGGAGGACGAGAGGCGCGCGGGTCCA 1856

QY 301 CCAGAACACACCCCGGGCTGTGCTGGTCCCGAGTAGCCTCCGGATTGGAATCATTTCCCGC 360
DB 1857 CCAGAACACACCCCGGGCTGTGCTGGTCCCGAGTAGCCTCCGGATTGGAATCATTTCCCGC 1916

QY 361 AGGTTCAACGAGCTGCGTGTGTTACTCACCGTGGGACACGACGACGAGAAAACCTTCGGC 420
DB 1917 AGGTTCAACGAGCTGCGTGTGTTACTCACCGTGGGACACGACGACGAGAAAACCTTCGGC 1976

QY 421 GCTGCATATCGTTGTTGGGACTCGCTGCGCATGATGTCTCAGTCCACCAACACAG 480
DB 1977 GCTGCATATCGTTGTTGGGACTCGCTGCGCATGATGTCTCAGTCCACCAACACAG 2036

QY 481 CACACTCTTCGCTACTCCGTGTCCTGCTGGGCTACGGCTTCTACGGGGACATCAAA 540
DB 2037 CACACTCTTCGCTACTCCGTGTCCTGCTGGGCTACGGCTTCTACGGGGACATCAAA 2096

QY 541 GGACAGTGAAGAAACGGTGGTGGTCTTGGCCAGATACACATTTTCAGGTTTAAAGAC 600
DB 2097 GGACAGTGAAGAAACGGTGGTGGTCTTGGCCAGATACACATTTTCAGGTTTAAAGAC 2156

QY 601 CTTCTCTCCACCACTCTATGAAGGACAGTGTCTTCTCTCTGACCAACACACGGT 660
DB 2157 CTTCTCTCCACCACTCTATGAAGGACAGTGTCTTCTCTCTGACCAACACACGGT 2216

QY 661 GGGATCTCCAAAGGATAGGAAAGCCCTGCGGGCAGGATGCTTTGTTGAGGCAAGCAA 720
DB 2217 GGGATCTCCAAAGGATAGGAAAGCCCTGCGGGCAGGATGCTTTGTTGAGGCAAGCAA 2276

QY 721 GCAGAGCTGGAGGAGGACGAGAAAGCACTGTATGTTTGGAGCTGGGAGGACGT 780
DB 2277 GCAGAGCTGGAGGAGGACGAGAAAGCACTGTATGTTTGGAGCTGGGAGGACGT 2336

QY 781 GGAGAGTGGCAAGTCTGTGGAAGTCTTGGCCATCAATGCCACAAACATGTCCTG 840
DB 2337 GGAGAGTGGCAAGTCTGTGGAAGTCTTGGCCATCAATGCCACAAACATGTCCTG 2396

QY 841 TGCTTGTCCCGAGCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTC 900
DB 2397 TGCTTGTCCCGAGCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTC 2456

QY 901 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCA 960
DB 2457 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCA 2516

QY 961 CACCAACAGCAGGACCCAG 979
DB 2517 CACCAACAGCAGGACCCAG 2535
```

```
RESULT 10
ABA96945
ID ABA96945 standard; cDNA; 4463 BP.
XX
AC ABA96945;
XX
DT 20-MAY-2002 (first entry)
XX
DE Human ceramide kinase hCERK1-encoding cDNA.
XX
KW Human; ceramide kinase; hCERK1; drug screening; gene therapy;
KW neurological disease; inflammation; human immunodeficiency virus;
KW HIV infection; type 2 diabetes; obesity; sepsis; arteriosclerosis;
KW cancer; neuroprotective; antiinflammatory; anti-HIV; antidiabetic;
KW anorectic; antibacterial; antiseptic; antiarteriosclerotic; cytostatic;
KW gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 124..1737
FT /*tag= a
FT /product= "Human ceramide kinase hCERK1"
XX
PN WO200196575-A1.
XX
PD 20-DEC-2001.
XX
PF 11-JUN-2001; 2001WO-JP004889.
XX
PR 14-JUN-2000; 2000JP-00178039.
XX
(SANY ) SANKYO CO LTD.
XX
PI Sugiura M, Kono K, Kohama T;
XX
WPI; 2002-179513/23.
DR P-PSDB; AAM49115.
XX
Human ceramide kinase gene and the enzyme encoded by it for screening
substances as drugs for neurological, inflammatory and other disorders.
XX
Claim 5; Page 46-53; 61pp; Japanese.
XX
This sequence represents cDNA encoding a human ceramide kinase designated
hCERK1. The invention relates to hCERK1, nucleic acids encoding it,
expression vectors and host cells containing hCERK1 nucleic acids, the
recombinant production of hCERK1 and antibodies specific for hCERK1. The
invention also encompasses methods of isolating hCERK1 from samples, the
use of hCERK1 in drug screening, and the use of hCERK1 nucleic acid
sequences in gene therapy. hCERK1 mediates the ATP-dependent 1-
phosphorylation of ceramides and can be used to screen for therapeutic
and preventive agents for a wide range of disorders. Such disorders
include neurological disease, inflammation, human immunodeficiency virus
(HIV) infection, type 2 diabetes, obesity, sepsis, arteriosclerosis and
cancer
XX
SQ Sequence 4463 BP; 1026 A; 1135 C; 1217 G; 1084 T; 0 U; 1 Other;
Query Match 91.7%; Score 898.2; DB 6; Length 4463;
Best Local Similarity 96.3%; Pred. No. 6.8e-257;
Matches 943; Conservative 0; Mismatches 3; Indels 33; Gaps 1;

QY 1 ACCAAGCATTTACTGTTATTTATCAACCCGTTTGGAGGAAAAGACAAAGCAAGCGGAT 60
DB 510 ACCAAGCATTTACTGTTATTTATCAACCCGTTTGGAGGAAAAGACAAAGCAAGCGGAT 569

QY 61 ATATGAAAGAAAAGTGGCACCACCTGTTTCACTTGTAGCTTCCATCACCCTGACATCATCGG 120
DB 570 ATATGAAAGAAAAGTGGCACCACCTGTTTCACTTGTAGCTTCCATCACCCTGACATCATCG- 628

QY 121 TAACAAATTCATGTTAACTATGTAGAAAGTAATTACTGAACATGCTAATCAGGCCAAGGA 180
```

Db 629 -----TTACTGACATGCTATATCAGGCCAAGA 656  
Qy 181 GACTCTGTATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGTCGGCGAGATGG 240  
Db 657 GACTCTGTATGAGATTAAATAGACAAATACGACGGCATCGTCTGTGTCGGCGAGATGG 716  
Qy 241 TATGTTTACGAGGTTCTGTCAGGTTCTGATTTGGGAGGACGACAGAGGCGCGGGGTGA 300  
Db 717 TATGTTTACGAGGTTCTGTCAGGTTCTGATTTGGGAGGACGACAGAGGCGCGGGGTGA 776  
Qy 301 CCAGAACACACCCCGGGCTGTGCTGCTCCCAAGTAGCTCCGGAATTGGAATCATTTCCCGC 360  
Db 777 CCAGAACACACCCCGGGCTGTGCTGCTCCCAAGTAGCTCCGGAATTGGAATCATTTCCCGC 836  
Qy 361 AGGGTCAACGGACTCGTGTGTTACTTCCACCGTGGGCAACGACGACGAGAAACCTCGGC 420  
Db 837 AGGTCCAAACGGACTCGTGTGTTACTTCCACCGTGGGCAACGACGACGAGAAACCTCGGC 896  
Qy 421 GCTGCATATCGTGTGTTGGGACTCGCTGGCCATGATGTCTCAGTCCACCACAAACAG 480  
Db 897 GCTGCATATCGTGTGTTGGGACTCGCTGGCCATGATGTCTCAGTCCACCACAAACAG 956  
Qy 481 CACACTCTCTTCTGCTACTCTCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCA 540  
Db 957 CACACTCTCTTCTGCTACTCTCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCA 1016  
Qy 541 GGACAGTGAAGAAACGGTGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGAC 600  
Db 1017 GGACAGTGAAGAAACGGTGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGAC 1076  
Qy 601 TTCTCTCTCCACCATCTGCTATGAAGGACAGTGTCTTCTCTCTGCAACAACACCGGT 660  
Db 1077 TTCTCTCTCCACCATCTGCTATGAAGGACAGTGTCTTCTCTCTGCAACAACACCGGT 1136  
Qy 661 GGGATCTCAAGGGATAGAAAGCCCTGCGGCGAGATGCTTTGTCAGGCAAGCA 720  
Db 1137 GGGATCTCAAGGGATAGAAAGCCCTGCGGCGAGATGCTTTGTCAGGCAAGCA 1196  
Qy 721 GCAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGGAGCTGCGGAGGAGCT 780  
Db 1197 GCAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGGAGCTGCGGAGGAGCT 1256  
Qy 781 GGAGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGGCCATCAATGCCAACAACATGCTCTG 840  
Db 1257 GGAGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGGCCATCAATGCCAACAACATGCTCTG 1316  
Qy 841 TGCTTTGTCGCGGAGCCCGAGGGCTCTCCCGGCTGCCACTTGGGAGAGCGGCTTTC 900  
Db 1317 TGCTTTGTCGCGGAGCCCGAGGGCTCTCCCGGCTGCCACTTGGGAGAGCGGCTTTC 1376  
Qy 901 TGACCTCATCTCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCA 960  
Db 1377 TGACCTCATCTCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCA 1436  
Qy 961 CACCAACACGAGGACCAG 979  
Db 1437 CACCAACACGAGGACCAG 1455

RESULT 11  
ADN62844  
ID ADN62844 standard; DNA; 1740 BP.  
XX  
AC ADN62844;  
XX  
DT 01-JUL-2004 (first entry)  
XX  
DE Human NOV9a DNA.  
XX  
ds; gene; human; NOVX; metabolic disorder; diabetes; obesity;  
KW infectious disease; anorexia; cancer; cancer-associated cachexia;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW immune disorder; haematopoietic disorder; dyslipidaemia;

KW metabolic syndrome X; wasting disorder.  
XX  
OS Homo sapiens.  
XX  
PN US2004038223-A1.  
XX  
PD 26-FEB-2004.  
XX  
PF 01-OCT-2002; 2002US-00262511.  
XX  
PR 02-OCT-2001; 2001US-0326483P.  
PR 05-OCT-2001; 2001US-0327435P.  
PR 05-OCT-2001; 2001US-0327449P.  
PR 09-OCT-2001; 2001US-0327917P.  
PR 09-OCT-2001; 2001US-0328029P.  
PR 09-OCT-2001; 2001US-0328044P.  
PR 09-OCT-2001; 2001US-0328056P.  
PR 12-OCT-2001; 2001US-0328849P.  
PR 15-OCT-2001; 2001US-0329414P.  
PR 17-OCT-2001; 2001US-0330142P.  
PR 18-OCT-2001; 2001US-0330309P.  
PR 22-OCT-2001; 2001US-0341058P.  
PR 24-OCT-2001; 2001US-0339266P.  
PR 29-OCT-2001; 2001US-0343629P.  
PR 29-OCT-2001; 2001US-0349575P.  
PR 01-NOV-2001; 2001US-0346357P.  
PR 17-APR-2002; 2002US-0373260P.  
PR 19-APR-2002; 2002US-0373815P.  
PR 19-APR-2002; 2002US-0373817P.  
PR 19-APR-2002; 2002US-0373826P.  
PR 19-APR-2002; 2002US-0373884P.  
PR 22-APR-2002; 2002US-0374977P.  
PR 16-MAY-2002; 2002US-0381037P.  
PR 16-MAY-2002; 2002US-0381038P.  
PR 16-MAY-2002; 2002US-0381042P.  
PR 17-MAY-2002; 2002US-0381642P.  
PR 28-MAY-2002; 2002US-0383656P.  
PR 29-MAY-2002; 2002US-0383831P.  
PR 25-JUN-2002; 2002US-0391335P.  
XX  
PA (SMIT/) SMITHSON G.  
PA (MILL/) MILLET I.  
PA (PEYM/) PEYMAN J A.  
PA (KEKU/) KEKUDA R.  
PA (JUJ/) JU J.  
PA (LILL/) LI L.  
PA (GUOX/) GUO X.  
PA (PATY/) PATTURAJAN M.  
PA (SPYT/) SPYTEK K A.  
PA (EDIN/) EDINGER S R.  
PA (ELLE/) ELLERMAN K.  
PA (MALY/) MALYANKAR U M.  
PA (ORTT/) ORT T.  
PA (GORM/) GORMAN L.  
PA (ZERH/) ZERHUSEN B D.  
PA (ANDE/) ANDERSON D W.  
PA (ZHON/) ZHONG M.  
PA (CATT/) CATTERTON E.  
PA (JIW/) JI W.  
PA (MILL/) MILLER C E.  
PA (RAST/) RASTELLI L.  
PA (STON/) STONE D J.  
PA (PENA/) PENNA C E A.  
PA (SHEN/) SHENOY S G.  
PA (SHIM/) SHIMKETS R A.  
PA (ROTH/) ROTHENBERG M E.  
PA (LEAC/) LEACH M D.  
PA (AGEE/) AGEE M L.  
PA (BERG/) BERGHS C.  
PA (DIPI/) DIPIPO V A.  
PA (EISE/) EISEN A.  
PA (GANG/) GANGOLLI E A.  
PA (RIBG/) RIEGER D K.

PA (SPAD/) SPADERNA S K.

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;

PI Paturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;

PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterson E;

PI Ji W, Miller CE, Rasceili L, Stone D, Pena CE, Shenoy SG;

PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;

PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;

XX WPI: 2004-213931/20.

DR P-PSDB; ADN62845.

XX Isolated NOVX polypeptides and nucleic acids, useful for preventing,

PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.

XX Claim 20; SEQ ID NO 39; 395pp; English.

XX The invention relates to isolated NOVX polypeptides and polynucleotides.

CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or

CC treat a medical condition in human related to the aberrant expression and

CC activity of NOVX polypeptides. For example, NOVX polypeptides and

CC polynucleotides may be used to treat disorders associated with decreased

CC expression or activity of NOVX by supplementing the patient our

CC production or to rectify mutations. Conversely, antisense NA molecules

CC may be administered to down regulate expression of NOVX polypeptides by

CC binding with the cells own genes and preventing their expression. NOVX

CC polynucleotides and complementary sequences may also be used as DNA

CC probes in diagnostic assays to detect and quantitate the presence of

CC similar sequences in samples, and so which patients may be in need of

CC restorative therapy. NOVX polypeptides may also be used as antigens in

CC the production of antibodies and in assays to identify modulators

CC (agonists and antagonists) of the expression and activity of NOVX. The

CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be

CC used to modulate NOVX polynucleotide expression and activity of NOVX

CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as

CC diagnostic agents for detecting the presence of NOVX in samples. NOVX

CC polypeptides and polynucleotides may be used in this way to prevent,

CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious

CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative

CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,

CC haematopoietic disorders, and the various dyslipidaemias, metabolic

CC disturbances associated with obesity, the metabolic syndrome X and

CC wasting disorders associated with chronic diseases and various cancers.

CC They may also be used as antibacterial agents. The present sequence

CC represents DNA encoding a human NOVX protein.

XX Sequence 1740 BP; 388 A; 477 C; 530 G; 345 T; 0 U; 0 Other;

SQ

Query Match 89.7%; Score 878.4; DB 12; Length 1740;

Best Local Similarity 96.3%; Pred. No. 3.4e-251;

Matches 944; Conservative 0; Mismatches 1; Indels 35; Gaps 3;

QY 1 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGGAAAGACAAGCAAGCGGAT 60

DB 462 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGGAAAGACAAGCAAGCGGAT 521

QY 61 ATATGAAAGAAAGTGGACCACTGTTTCACTTAGCCTCCATCCACCTGACATCATCGG 120

DB 522 ATATGAAAGAAAGTGGACCACTGTTTCACTTAGCCTCCATCCACCTGACATCATCG- 580

QY 121 TAACAAATTTCTATGTTAACTATGTAGAAGTAAATTAAGTAAATTAAGTAAATTAAGTAA 180

DB 581 -----TTACTGAACATGCTATCATCAGGCCAAGGA 608

QY 181 GACTCTGTATGAGATTACATAGACAATAACAGCGGATTCGTGTGTGCGGGAGAT-G 239

DB 609 GACTCTGTATGAGATTACATAGACAATAACAGCGGAT-GTCTGTGTGCGGGAGATCG 667

QY 240 GTATGTTTCAGCGAGTGTGCAACGCTGTTGTTGGAGGACCGACAGGAGCCCGGGTCG 299

DB 668 GTATGTTTCAGCGAGTGTGCAACGCTGTTGTTGGAGGACCGACAGGAGCCCGGGTCG 727

QY 300 ACCAGAACCCCGCGGCTGTGTGTTCCCGAGTAGCCTCCGGATTGGAATCATTCGCCG 359

DB 728 ACCAGAACCCCGCGGCTGTGTGTTCCCGAGTAGCCTCCGGATTGGAATCATTCGCCG 787

QY 360 CAGGCTCAACCGAGTGGTGTGTACTTCCACCGTGGGCACAGCGACGCAAAACCTTCGG 419

DB 788 CAGGCTCAACCGAGTGGTGTGTACTTCCACCGTGGGCACAGCGACGCAAAACCTTCGG 847

QY 420 CGCTGCATATCGTTGTTGGGGACTCGCTGGCCATGGATGTCTCTAGTCCACCAACA 479

DB 848 CGCTGCATATCGTTGTTGGGGACTCGCTGGCCATGGATGTCTCTAGTCCACCAACA 907

QY 480 GCACATCTCTTCCGTACTCTCGTGTCCGTGTGGGCTACGGCTTACGGGACATCATCA 539

DB 908 GCACATCTCTTCCGTACTCTCGTGTCCGTGTGGGCTACGGCTTACGGGACATCATCA 967

QY 540 AGGACAGTGAAGAAACCGTGTGGGTCTGCCAGATACGACTTTTTCAGGTTTAAAGA 599

DB 968 AGGACAGTGAAGAAACCGTGTGGGTCTGCCAGATACGACTTTTTCAGGTTTAAAGA 1027

QY 600 CCTTCTCTCCCACTGCTATGAAGGACAGTGTCTTCTCTCTGCAACACACACGG 659

DB 1028 CCTTCTCTCCCACTGCTATGAAGGACAGTGTCTTCTCTCTGCAACACACACGG 1087

QY 660 TGGGATCTCAAGGATAGAAAGCCCTGCGGGCAGGATGCTTTTTCAGGCAAAAGCA 719

DB 1088 TGGGATCTCAAGGATAGAAAGCCCTGCGGGCAGGATGCTTTTTCAGGCAAAAGCA 1147

QY 720 AGCAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGGAAAGCTTGGAAAGCTTGGAGGACG 779

DB 1148 AGCAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGGAAAGCTTGGAAAGCTTGGAGGACG 1207

QY 780 TGGAGGAGTGGCAAGTCTGTGTGGGAAAGTCTTGGCCATCAATGCCACAAACATGTCTCT 839

DB 1208 TGGAGGAGTGGCAAGTCTGTGTGGGAAAGTCTTGGCCATCAATGCCACAAACATGTCTCT 1267

QY 840 GTGCTTGTGCGCGGAGCCCGAGGGGCTCTCCCGGCTGCCCTTGGGAGACGGGTCTT 899

DB 1268 GTGCTTGTGCGCGGAGCCCGAGGGGCTCTCCCGGCTGCCCTTGGGAGACGGGTCTT 1327

QY 900 CTGACCTCATCTCCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGC 959

DB 1328 CTGACCTCATCTCCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGC 1387

QY 960 ACACCAACCGACGAGCAG 979

DB 1388 ACACCAACCGACGAGCAG 1407

RESULT 12

AAS77728

ID AAS77728 standard; cDNA; 2241 BP.

XX AAS77728;

AC AAS77728;

XX 13-FEB-2002 (first entry)

DT

XX DNA encoding novel human diagnostic protein #13532.

DE

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

PN

XX 11-OCT-2001.

PD

XX 30-MAR-2001; 2001WO-US008631.

PF

XX 31-MAR-2000; 2000US-00540217.

PR

XX 23-AUG-2000; 2000US-00649167.

FR

XX (HYSE-) HYSEQ INC.

XX Dmanac RT, Liu C, Tang YT;  
XX WPI; 2001-639362/73.  
DR P-PSDB; ABG13541.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
PS Claim 1; SEQ ID NO 13532; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 2241 BP; 453 A; 627 C; 700 G; 461 T; 0 U; 0 Other;

Query Match 63.6%; Score 622.2; DB 5; Length 2241;  
Best Local Similarity 81.6%; Pred. No. 1.3e-174;  
Matches 827; Conservative 0; Mismatches 3; Indels 183; Gaps 2;

Qy 150 TAATTAATGACATGCTTAATCAGGCAAGGAGACTCTGTATGAGATTAAACATAGACAAAT 209  
Db TGAATTACTGAACATGCTTAATCAGGCAAGGAGACTCTGTATGAGATTAAACATAGACAAAT 955  
Qy 210 ACGAGCGCATGCTGTCTGCGCGGAGATGTTTCCAGCGAGTGCTGCACGCTCTGA 269  
Db ACGAGCGCATGCTGTCTGCGCGGAGATGTTTCCAGCGAGTGCTGCACGCTCTGA 1015  
Qy 270 TTGGGAGGACGACGAGGAGCGCGGGGTCGACCAAGAACACCCCGGGGCTGTGCTGTCC 329  
Db TTGGGAGGACGACGAGGAGCGCGGGGTCGACCAAGAACACCCCGGGGCTGTGCTGTCC 1075  
Qy 330 CCAGTAGCTCGGATTGGAATCATCTCCGCA----- 361  
Db CCAGTAGCTCGGATTGGAATCATCTCCGCAAGAGCATGTGGGCCCCGGTGAGAACGCTG 1135  
Qy 362 -----GGGTCAACGAGCTCGCTGTGTACTCCACCGTGGGACCAAGCAGCGAG 410  
Db GTGGCTTGAGCGGGTCAACGAGCTCGCTGTGTACTCCACCGTGGGACCAAGCAGCGAG 1195  
Qy 411 AAACCTCGGCGCTGCATATCGTTGTT----- 436  
Db AAACCTCGGCGCTGCATATCGTTGTTGCTGTGCTGCTGCGGCTCGGAACACCGCGCAT 1255  
Qy 437 ----- 436  
Db CCGCGCATACTGCTGTGGCAGTGGGCACGCGGAGCTGTGCTGGGCTGCCAGCTGTGA 1315  
Qy 437 -----GGGAGCTCGC 446  
Db ACGCATGCTCTGTGAGGCTCGAGGCTTCAGTCCAGGATCGAGAGCCCCGGGAGCTCGC 1375

Qy 447 TGCCATGGATGTGTCCTCAGTCCACCAACAGCACATCCTTCGCTACTCCTGTGCCC 506  
Db TGCCATGGATGTGTCCTCAGTCCACCAACAGCACATCCTTCGCTACTCCTGTGCCC 1435  
Qy 507 TGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAAACGGTGTGG 566  
Db TGCTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAAACGGTGTGG 1495  
Qy 567 GTCTTGCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACACTGTCTATGAAG 626  
Db GTCTTGCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACACTGTCTATGAAG 1555  
Qy 627 GGACAGTGTCTCTCTCCCTGCAACACACAGGTGGGATCTCCAGGATAGGAGCCCT 686  
Db GGACAGTGTCTCTCTCCCTGCAACACACAGGTGGGATCTCCAGGATAGGAGCCCT 1615  
Qy 687 GCCGGGACAGATGCTTTGTTTTCAGGCAAGGAGCAGTGGAGGAGGAGCAGAGA 746  
Db GCCGGGACAGATGCTTTGTTTTCAGGCAAGGAGCAGTGGAGGAGGAGGAGCAGAGA 1675  
Qy 747 AAGCACTGTATGTTTGGAAAGCTGCGGAGGACGTGGAGGAGTGGCAAGTCTGTGGGA 806  
Db AAGCACTGTATGTTTGGAAAGCTGCGGAGGACGTGGAGGAGTGGCAAGTCTGTGGGA 1735  
Qy 807 AGTTTCTGGCCATCAATGCCACAAACATGTCCTGTGCTGTGCGCGGAGCCCCAGGGGCC 866  
Db AGTTTCTGGCCATCAATGCCACAAACATGTCCTGTGCTGTGCGCGGAGCCCCAGGGGCC 1795  
Qy 867 TCTCCCGGCTGCCACTTGGGAGACGGGTCTTCTGACCTCATCTCCGGAATGCT 926  
Db TCTCCCGGCTGCCACTTGGGAGACGGGTCTTCTGACCTCATCTCCGGAATGCT 1855  
Qy 927 CCAGGTTCAATTTTCTGAGATTCTTCATCAGGCACACCAACAGCAGGACCCAG 979  
Db CCAGGTTCAATTTTCTGAGATTCTTCATCAGGCACACCAACAGCAGGACCCAG 1908

RESULT 13  
ADS10370  
ID ADS10370 standard; DNA; 4702 BP.  
XX  
AC ADS10370;  
XX  
DT 16-DEC-2004 (first entry)  
XX  
DE Human therapeutic DNA - SEQ ID 607.  
XX  
KW antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnery;  
KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;  
KW aplastic anaemia; cancer; wound healing; gene therapy; ds; gene.  
XX  
OS Homo sapiens.  
XX  
PN WO2004080148-A2.  
XX  
PD 23-SEP-2004.  
XX  
PF 30-SEP-2003; 2003WO-US030720.  
XX  
PR 02-OCT-2002; 2002US-0416186P.  
XX  
PA (NUVE-) NUVELO INC.  
XX  
PI Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y;  
PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;  
XX  
XX WPI; 2004-669857/65.  
DR P-PSDB; ADS11054.  
XX  
PT New polynucleotide, useful in preparing a composition for diagnosing or  
PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,  
PT aplastic anemia or cancer for promoting wound healing.  
XX

PS Claim 1; SEQ ID NO 607; 718pp; English.

XX The invention relates to a novel isolated polynucleotide and the encoded  
XX polypeptide. The molecules of the invention demonstrate antiinflammatory,  
CC neuroprotective, antianemic, cytostatic and vulnerary activities and may  
CC be useful in preparing a composition for diagnosing or treating  
CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell  
CC disorders, such as aplastic anaemia or cancer, as well as for promoting  
CC wound healing. The molecules may also be utilised during gene therapy  
CC procedures. The current sequence is that of a human therapeutic DNA of  
CC the invention. The current sequence is not shown explicitly within the  
CC specification but can be accessed from the WIPO web-site.

XX  
SQ Sequence 4702 BP; 970 A; 1226 C; 1364 G; 1118 T; 0 U; 24 Other;

Query Match 63.1%; Score 618; DB 13; Length 4702;  
Best Local Similarity 100.0%; Pred. No. 3.4e-173;  
Matches 618; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 362 GGGTCAACGGACTCGGTGTGTACTCCACCGTGGGACACGACGAGAAACCTCGGCG 421  
Db |||||||  
Qy 2044 GGGTCAACGGACTCGGTGTGTACTCCACCGTGGGACACGAGAAACCTCGGCG 2103  
Db |||||||  
Qy 422 CTGCATATCGTGTGGGACTCGCTGGCCATGATGTCCTCAGTCCACCAACAGC 481  
Db |||||||  
Qy 2104 CTGCATATCGTGTGGGACTCGCTGGCCATGATGTCCTCAGTCCACCAACAGC 2163  
Db |||||||  
Qy 482 ACCTCTCTCGTACTCGGTGTCTCGTGGCTTCTACGGGGACATCATCAAG 541  
Db |||||||  
Qy 2164 ACCTCTCTCGTACTCGGTGTCTCGTGGCTTCTACGGGGACATCATCAAG 2223  
Db |||||||  
Qy 542 GACAGTGAAGAAACCGTGTGGGTCTGGCCAGATACGACTTTTCAGGTTTAAAGACC 601  
Db |||||||  
Qy 2224 GACAGTGAAGAAACCGTGTGGGTCTGGCCAGATACGACTTTTCAGGTTTAAAGACC 2283  
Db |||||||  
Qy 602 TTCTCTCCACCACTGCTATGAAGGACAGTGTCCTTCCTCCCTGCACAAACACCGGTG 661  
Db |||||||  
Qy 2284 TTCTCTCCACCACTGCTATGAAGGACAGTGTCCTTCCTCCCTGCACAAACACCGGTG 2343  
Db |||||||  
Qy 662 GGATCTCCAGGGATAGGAGCCCTCGCGGACAGGATCTTGTTCAGGCAAGCAAG 721  
Db |||||||  
Qy 2344 GGATCTCCAGGGATAGGAGCCCTCGCGGACAGGATCTTGTTCAGGCAAGCAAG 2403  
Db |||||||  
Qy 722 CAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGAAGCTGCGGAGGACGTG 781  
Db |||||||  
Qy 2404 CAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGAAGCTGCGGAGGACGTG 2463  
Db |||||||  
Qy 782 GAGGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCAAAACATGTCTCTGT 841  
Db |||||||  
Qy 2464 GAGGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCAAAACATGTCTCTGT 2523  
Db |||||||  
Qy 842 GCTTGTCCCGGAGCCCGGAGGCTCTCCCGGCTGCCACTTGGAGAGCGGTCTTCT 901  
Db |||||||  
Qy 2524 GCTTGTCCCGGAGCCCGGAGGCTCTCCCGGCTGCCACTTGGAGAGCGGTCTTCT 2583  
Db |||||||  
Qy 902 GACCTCATCTCATCCGAAATGCTCCAGGTTCAATTTCTGAGATTTCATCAGGCAC 961  
Db |||||||  
Qy 2584 GACCTCATCTCATCCGAAATGCTCCAGGTTCAATTTCTGAGATTTCATCAGGCAC 2643  
Db |||||||  
Qy 962 ACCAACGACGAGGACAC 979  
Db |||||||  
Qy 2644 ACCAACGACGAGGACAC 2661  
Db |||||||

RESULT 14  
AAC76031  
ID AAC76031 standard; cDNA; 753 BP.  
XX  
AC AAC76031;  
XX  
DT 08-FEB-2001 (first entry)  
XX  
DE Human ORFX ORF1586 polynucleotide sequence SEQ ID NO:3171.

Query Match 61.2%; Score 598.8; DB 3; Length 753;  
Best Local Similarity 99.7%; Pred. No. 7e-168;  
Matches 600; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 360 CAGGTCACGGACTCGGTGTGTACTCCACCGTGGGACACGAGAAACCTCGG 419  
Db |||||||  
Qy 131 CAGGTCACGGACTCGGTGTGTACTCCACCGTGGGACACGAGAAACCTCGG 190  
Db |||||||

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
KW vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;  
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
KW cholesterol ester storage; systemic lupus erythematosus; infection;  
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
KW thrombosis; contraceptive; ss.

XX Homo sapiens.  
OS  
XX  
XX WO200058473-A2.  
PN  
XX  
PD 05-OCT-2000.  
XX  
PF 31-MAR-2000; 2000WO-US008621.  
XX  
XX 31-MAR-1999; 99US-0127607P.  
PR 02-APR-1999; 99US-0127636P.  
PR 05-APR-1999; 99US-0127728P.  
PR 30-MAR-2000; 2000US-00540763.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX Shimkets RA, Leach M;  
PI  
XX  
XX WPI; 2000-602362/57.  
DR  
XX P-PSDB; AAB41822.  
XX  
XX Novel nucleic acids and peptides derived from open reading frame X,  
PT useful for treating e.g. cancers, proliferative disorders,  
PT neurodegenerative disorders and cardiovascular disease.  
XX  
PS Claim 5; Page 2392; 5507pp; English.  
XX  
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;  
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;  
CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;  
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;  
CC antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The  
CC sequences can be used for determining the presence of or predisposition  
CC to, or preventing or treating pathological conditions associated with an  
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX  
CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
CC used to treat cancers, proliferative disorders, neurodegenerative  
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
XX  
SQ Sequence 753 BP; 165 A; 195 C; 207 G; 186 T; 0 U; 0 Other;

```
QY 420 CGCTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTGCTCAGTCACCAACA 479
Db |||||
191 CGCTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTGCTCAGTCACCAACA 250
QY 480 GCACACTCTTCGCTACTCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCA 539
Db |||||
251 GCACACTCTTCGCTACTCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGACATCATCA 310
QY 540 AGGACAGTGAAGAAACCGTGGTGGTCTTTGCCAGATACGACTTTTCAGGTTTAAAGA 599
Db |||||
311 AGGACAGTGAAGAAACCGTGGTGGTCTTTGCCAGATACGACTTTTCAGGTTTAAAGA 370
QY 600 CTTTCTCTCCCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACACACAGG 659
Db |||||
371 CTTTCTCTCCCACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACACACAGG 430
QY 660 TGGGATCTTCAAGGATAGGAAGCCCTCGCGGACAGGATGCTTTGTTGACGCAAGCA 719
Db |||||
431 TGGGATCTTCAAGGATAGGAAGCCCTCGCGGACAGGATGCTTTGTTGACGCAAGCA 490
QY 720 AGCAGAGCTGGAGGAGCAGAGAAAGCACTGTATGTTTGAAGCTGCGGAGGACG 779
Db |||||
491 AGCAGAGCTGGAGGAGCAGAGAAAGCACTGTATGTTTGAAGCTGCGGAGGATG 550
QY 780 TGGAGAGTGGCAAGTCTGTGGGAAGTTTCTGCCATCAATGCCACAAACATGTCCT 839
Db |||||
551 TGGAGAGTGGCAAGTCTGTGGGAAGTTTCTGCCATCAATGCCACAAACATGTCCT 610
QY 840 GTGCTTGTCCGCGAGCCCGAGGGCTCTCTCCCGGCTGCCACATGGGAGACGGTCTT 899
Db |||||
611 GTGCTTGTCCGCGAGCCCGAGGGCTCTCTCCCGGCTGCCACATGGGAGACGGTCTT 670
QY 900 CTGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGC 959
Db |||||
671 CTGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCTGCG 730
QY 960 AC 961
Db ||
731 AC 732

RESULT 15
ID AAS77731 standard; cDNA; 1570 BP.
AC
XX
XX AAS77731;
DE 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #13535.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX P-PSDB; ABG13544.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
```

```
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 13535; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful for medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 1570 BP; 349 A; 446 C; 456 G; 319 T; 0 U; 0 Other;

Query Match 55.5%; Score 543; DB 5; Length 1570;
Best Local Similarity 100.0%; Pred. No. 4.8e-151;
Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 GGGGACTCGCTGGCCATGGATGTGCTCCTCAGTCACCAACAGCAGACTCTTCGCTAC 496
Db |||||
595 GGGGACTCGCTGGCCATGGATGTGCTCCTCAGTCACCAACAGCAGACTCTTCGCTAC 754
QY 497 TCCGTGTCCCTGTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGAGAAGAA 556
Db |||||
755 TCCGTGTCCCTGTGGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGAGAAGAA 814
QY 557 CGGTGGTGGTCTTGCAGATACGACTTTTCAGGTTTAAAGACTTCTCTCCACCAC 616
Db |||||
815 CGGTGGTGGTCTTGCAGATACGACTTTTCAGGTTTAAAGACTTCTCTCCACCAC 874
QY 617 TGCTATGAAGGACAGTGTCTCTCCCTGCACAAACACACGCTGGGATCTCCAAGGAT 676
Db |||||
875 TGCTATGAAGGACAGTGTCTCTCTCCCTGCACAAACACACGCTGGGATCTCCAAGGAT 934
QY 677 AGGAAGCCCTGCCGGGAGGATGCTTTGTTTGCAGGCAAGCAAGCAGCAGCTGGAGGAG 736
Db |||||
935 AGGAAGCCCTGCCGGGAGGATGCTTTGTTTGCAGGCAAGCAAGCAGCAGCTGGAGGAG 994
QY 737 GAGCAGAAGAAAGCACTGTATGTTTGGAAAGCTCGGAGGACGCTGGAGGAGTGCGAAGTC 796
Db |||||
995 GAGCAGAAGAAAGCACTGTATGTTTGGAAAGCTCGGAGGACGCTGGAGGAGTGCGAAGTC 1054
QY 797 GTCTGTGGGAGTGTCTGGCCATCAATGCCACAAACATGTCTGTGCTTGTGCGGGAGC 856
Db |||||
1055 GTCTGTGGGAGTGTCTGGCCATCAATGCCACAAACATGTCTGTGCTTGTGCGGGAGC 1114
QY 857 CCCAGGGGCTCTCTCCCGGCTGCCACATTTGGGAGACGGGTCTTCTGACCTCATCTCATC 916
Db |||||
1115 CCCAGGGGCTCTCTCCCGGCTGCCACATTTGGGAGACGGGTCTTCTGACCTCATCTCATC 1174
QY 917 CGGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATAGGACACACCAACAGAGGAC 976
Db |||||
1175 CGGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATAGGACACACCAACAGAGGAC 1234
QY 977 CAG 979
Db |||
1235 CAG 1237
```

Search completed: September 5, 2005, 20:25:07  
Job time : 578.067 secs

---



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2005, 09:12:32 ; Search time 179.716 Seconds  
(without alignments)  
8913.583 Million cell updates/sec

Title: US-10-631-958-1

Perfect score: 979

Sequence: 1 accaaagcattacttggtat.....acaccaaccagcaggaccag 979

Scoring table:

IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTUS COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	901.4	92.1	4432	4	US-09-774-528-148
2	78.6	8.0	1084	4	US-09-270-767-15155
3	77.7	7.9	2064	4	US-09-270-767-14306
4	69.8	7.1	498	4	US-09-893-737-59
5	55.5	5.6	901	4	US-09-270-767-30448
6	43.4	4.4	1050	4	US-09-252-991A-1269
7	43.4	4.4	1308	4	US-09-252-991A-1179
8	42.2	4.3	430	4	US-09-621-976-16656
9	40.8	4.2	1014	4	US-09-902-540-9500
10	40.8	4.2	14101	4	US-09-902-540-1080
11	40.4	4.1	1857	4	US-09-970-516-3
12	40.4	4.1	2380	4	US-09-817-676A-13
13	39.8	4.1	485	4	US-09-270-767-31476
14	39.4	4.0	2698	4	US-09-817-676A-11
15	38.8	4.0	1875	4	US-09-614-221A-399
16	38.4	3.9	2274	4	US-09-252-991A-3660
17	38.4	3.9	3147	4	US-09-902-540-3530
18	38.4	3.9	3297	4	US-09-252-991A-3615
19	38.4	3.9	18538	4	US-09-902-540-1169
20	37.8	3.9	395	4	US-09-894-844-45
21	37.8	3.9	4403765	3	US-09-103-840A-2
22	37.8	3.9	4411529	3	US-09-103-840A-1
23	37.2	3.8	6644	4	US-08-875-435B-5
24	36.8	3.8	601	4	US-09-949-016-62772
25	36.8	3.8	601	4	US-09-949-016-62773
26	36.8	3.8	103792	4	US-09-949-016-13553
27	36.8	3.8	103993	4	US-09-949-016-12317

## ALIGNMENTS

### RESULT 1

US-09-774-528-148

; Sequence 148, Application US/09774528

; Patent No. 6743619

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Zhou, Ping

; APPLICANT: Goodrich, Ryle

; APPLICANT: Liu, Chenghua

; APPLICANT: Asundi, Vinod

; APPLICANT: Ren, Feiyan

; APPLICANT: Zhang, Jie

; APPLICANT: Zhao, Qing A.

; APPLICANT: Yang, Yonghong

; APPLICANT: Xue, Aidong J.

; APPLICANT: Wehrman, Tom

; APPLICANT: Wang, Jian-Rui

; APPLICANT: Wang, Dunrui

; APPLICANT: Drmanac, Radoje T.

; TITLE OF INVENTION: No. 6743619el Nucleic Acids and

; FILE REFERENCE: Polyptides

; CURRENT APPLICATION NUMBER: US/09/774,528

; NUMBER OF SEQ ID NOS: 441

; SOFTWARE: pt\_FL\_genes Version 2.0

; SEQ ID NO 148

; LENGTH: 4432

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1129)..(2817)

; US-09-774-528-148

Query Match 92.1%; Score 901.4; DB 4; Length 4432;

Best Local Similarity 96.5%; Pred. No. 1.4e-266;

Matches 945; Conservative 0; Mismatches 1; Indels 33; Gaps 1;

Qy 1 ACCAAGCATTTACTGGTATTTATCAACCCGTTTGGAGGAAAGGCAAGGCAAGCGGAT 60

|||||

1590 ACCAAGCATTTACTGGTATTTATCAACCCGTTTGGAGGAAAGGCAAGGCAAGCGGAT 1649

Qy 61 ATATGAAGAAAGTGGGACCACTGTTCACCTTAGCTCCATCACCCTGACATCATCGG 120

|||||

1650 ATATGAAGAAAGTGGGACCACTGTTCACCTTAGCTCCATCACCCTGACATCATCG- 1708

Qy 121 TAACAAATTCATGTTAACTAGTAAAGTAACTTACTGAAATCTTAATCAATCAAGGCAAGGA 180

|||||

1709 -----TTACTGAACATGCTAATCAAGGCAAGGA 1736

Db

		; ORGANISM: Drosophila melanogaster	
		US-09-270-767-15155	
		Query Match 8.0%; Score 78.6; DB 4; Length 1084;	
		Best Local Similarity 49.1%; Pred. No. 1.5e-13;	
		Matches 238; Conservative 0; Mismatches 244; Indels 3; Gaps 1;	
Qy	150	TAATTACTGAACATGCTAAATCAGGCCAAAGGAGACTCTGTATGAGATTAAACATAGACAAAT	209
Db	926	TCACCACTCAGAGGGCAACCAAGTGAAGACATACCTTCTGAGCCATGATCTGGAGTAT	867
Qy	210	ACGAGCGCATGCTGTGTCCGGGAGATGGTATGTTTCAGCGAGGTCTGCACGGTCTGA	269
Db	866	ACGATGCGGTTTGTGTGTCCGAGCGATGCCACCGTAGCAGAGGTCAACAACGACTGA	807
Qy	270	TTGGGAGGACGACAGAGAGCGCCGGGGTCCAGCAACACCCCGGGCTGTGTGTCTCC	329
Db	806	TATTCCTCAAAATCCGAGAGTTTGGGACTGGACGAACAGCGGCCACCATATCAATTC	747
Qy	330	CCAGTAGCCTCCGGATTGGAAATCATCCCGCAGGGTCAACGGACTGGTGTGTACTCCA	389
Db	746	CGGT---CTGCCAGTGGTGTGATTCCTCCGTGGCAGCACCGACACCATTTGCGTATAGTA	690
Qy	390	CCGTGGCACACGACGACGCAAAACCTCGCGCTGCATATCGTTTGTGGGACTCGCTGG	449
Db	689	TGCACGGCACCGGGATGTGAGGACAGCGCTATCCATGTGATTCGGGCCAGCATCGGG	630
Qy	450	CCATGATGTCTCTCAGTCCACCAACAGCACACTCTTCCTCGTACTCCGTGTCTCCGTC	509
Db	629	GATTGGATGTGTGAGTGTGAGCAATGGCCAGTCCCTGCTCAGATTCTGTGCCAGTGTCC	570
Qy	510	TGGCTACGGCTTCTACGGGACATCATCAAGGACAGTGAGAGAAACGGTGGTGGGTG	569
Db	569	TGACTACGGGTACTCGGCGATGTGGCAGCCAGACGAGAGACTACCGCTGGATGGGAC	510
Qy	570	TTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGTCTATGAAGGA	629
Db	509	CGCGCGGTACGAGTACAGTGGCGTCAAGGCTTCTCTGAATATATCGCGGCTATGACCG	450
Qy	630	CAGTG 634	
Db	449	AACTG 445	
RESULT 3			
US-09-270-767-14306			
; Sequence 14306, Application US/09270767			
; Patent No. 6703491			
; GENERAL INFORMATION:			
; APPLICANT: Homburger et al.			
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster			
; FILE REFERENCE: File Reference: 7326-094			
; CURRENT APPLICATION NUMBER: US/09/270,767			
; CURRENT FILING DATE: 1999-03-17			
; NUMBER OF SEQ ID NOS: 62517			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 14306			
; LENGTH: 2064			
; TYPE: DNA			
; ORGANISM: Drosophila melanogaster			
US-09-270-767-14306			
		Query Match 7.9%; Score 77; DB 4; Length 2064;	
		Best Local Similarity 48.9%; Pred. No. 7.2e-13;	
		Matches 237; Conservative 0; Mismatches 245; Indels 3; Gaps 1;	
Qy	150	TAATTACTGAACATGCTAAATCAGGCCAAAGGAGACTCTGTATGAGATTAAACATAGACAAAT	209
Db	722	TCACCACTCAGAGGGCAACCAAGTGAAGACATACCTTCTGAGCCATGATCTGGAGTAT	781
Qy	210	ACGAGCGCATGCTGTGTCCGGGAGATGGTATGTTTCAGCGAGGTCTGCACGGTCTGA	269
Db	782	ACGATGCGGTTTGTGTGTCCGAGCGATGGCCCGTAGCAGAGGTCAACAACGACTGA	841

QY 270 TTGGAGGACGACAGAGAGCGCGGGGTGACACAGAACACCCCGGGGTGTGTGTGTC 329  
Db 842 TATTCGCTCAAAATGCGAGAGTTGGGACTGGACAGAAACAGCGGCCACCATACATTCCAAGAC 901  
QY 330 CCAGTAGCTCCGGATTGGAATCATTTCCGCGAGGGTCAACGGACTGCGTGTGTACTCCA 389  
Db 902 CGGCT---CTGCGAGTGGTGTGATTCCGCTGGGAGACCGACACCATTTGGGTATAGTA 958  
QY 390 CGGTGGGACACGACGACGAGAAACCTCGGCGCTGCATATCGTTGTTGGGACTCGCTGG 449  
Db 959 TGCACGGCAGCGCGATGTGAGGACAGCGGCTATCCATGTGATTCTGGCGACGATCGG 1018  
QY 450 CCATGGATGTGCTCAGTCCACCAACAGACACACTCTTTCGCTACTCCGTTCCGTTCC 509  
Db 1019 GATTGGATGTGTGCAAGTGTGAGCAATGGGCACTCCCTGCTCAGATTCTGTGCCAGTGTCC 1078  
QY 510 TGGGCTACGGCTTCTACGCGGACATCATCAAGGACAGTGAGAGAAACGGTGTGGTGC 569  
Db 1079 TGAGCTACGGTACCTGGCGGATGTGGAGCCAGAGGAGAACTACCGCTGGATGGGAC 1138  
QY 570 TTGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAAGGA 629  
Db 1139 CGCGCGGTACGAGTACAGTGGCGTCAAGGCTTCTGTAATATCGCGGCTATGACGCG 1198  
QY 630 CAGTG 634  
Db 1199 AACTG 1203

## RESULT 4

US-09-893-737-59/c  
; Sequence 59, Application US/09893737  
; Patent No. 6822082  
; GENERAL INFORMATION:  
; APPLICANT: Sheppard, Paul O.  
; APPLICANT: Presnell, Scott R.  
; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS  
; FILE REFERENCE: 00-41  
; CURRENT APPLICATION NUMBER: US/09/893,737  
; PRIORITY FILING DATE: 2001-06-28  
; PRIOR APPLICATION NUMBER: US 60/215,446  
; PRIOR FILING DATE: 2000-06-30  
; NUMBER OF SEQ ID NOS: 329  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 59  
; LENGTH: 498  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(498)  
US-09-893-737-59

Query Match 7.1%; Score 69.8; DB 4; Length 498;  
Best Local Similarity 91.4%; Pred. No. 4.8e-11;  
Matches 74; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 360 CAGGCTCAACGAGCTCGTGTGTACTCCACGCTGGGACCGACGACGAGAAACCTGG 419  
Db 460 CAGGCTCAACGAGCTCGTGTGTACTCCACGCTGGGACCGACGACGAGAAACCTGG 401  
QY 420 CGCTGCATATCGTTGTGGG 440  
Db 400 CGCTGCATATCGTTGTGGG 380

## RESULT 5

US-09-270-767-30448  
; Sequence 30448, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 30448  
; LENGTH: 901  
; TYPE: DNA  
; ORGANISM: *Drosophila melanogaster*  
US-09-270-767-30448

Query Match 5.6%; Score 55; DB 4; Length 901;  
Best Local Similarity 57.1%; Pred. No. 2.5e-06;  
Matches 100; Conservative 0; Mismatches 75; Indels 0; Gaps 0;  
QY 782 GAGGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCCCTGT 841  
Db 368 GATCAGTGAAGTTGTGCGGGGCAATTTCTTTATGATCTCGGGCGCAACATACCTGC 427  
QY 842 GTTGTGCGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGAGCGGTCTTCT 901  
Db 428 GCCTGCGCAGGAGTCCCAATGGCATCTCCGTTACAGTCATCTGGGTGATGTTGCCCTG 487  
QY 902 GACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCA 956  
Db 488 GACCTGATTTCTCGTGAAGAAAGACCTCACTGCTCAACACGTCGCTTTTCTGCTCA 542

## RESULT 6

US-09-252-991A-1269  
; Sequence 1269, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIORITY FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 1269  
; LENGTH: 1050  
; TYPE: DNA  
; ORGANISM: *Pseudomonas aeruginosa*  
US-09-252-991A-1269

Query Match 4.4%; Score 43.4; DB 4; Length 1050;  
Best Local Similarity 49.3%; Pred. No. 0.01; Mismatches 116; Indels 0; Gaps 0;  
Matches 113; Conservative 0; Mismatches 75; Indels 0; Gaps 0;  
QY 307 CCACCCCGGGCTGTGCTGTCCAGTAGCTCCCGATGGAATCATTCGCCGAGGTC 366  
Db 691 CCATCTATGTGGGTCTACGCCACACAGTCTGCTCTGATCGGGCCACTTTCGCGGCTT 750  
QY 367 AACGGACTGGTGTGTACTCCACGTTGGGACAGCGACGAGAAACCTCGGCGCTGCA 426  
Db 751 CACCTATTTCTGCGCGATCTCCACGAGGTGAGCGGCTTCTCCGCTCGACCGTACCGCT 810  
QY 427 TATCGTTTGGGACTCGCTGGCCATGATGTCTCTCAGTCACACACACACACACT 486  
Db 811 GCTGCTGGTGTCTACGGGCTTGGGACGCTGTTGGGCAACATCTGCGGCCCTGGC 870  
QY 487 CTTTCGCTACTCCGTGTCCCTGTGGGTACGGCTTCTACGGGACATC 535  
Db 871 CGACCGCATACCATCGGGTCTCGGCTTTCGGGCTGCTGGCGGCAATC 919

## RESULT 7

```
US-09-252-991A-1179
; Sequence 1179, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1179
; LENGTH: 1308
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1179

Query Match          4.4%; Score 43.4; DB 4; Length 1308;
Best Local Similarity 49.3%; Pred. No. 0.012;
Matches 113; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

Qy 307 CCACCCCGGCTGTGCTGTCGCCAGTAGCCTCCGGATTGGAATCATTCGCCGAGGTC 366
Db 741 CCACTATGTGGGTCTACGCACACAGTCTGCTGATCGCGCCACCTTCGCGGCTT 800

Qy 367 AACGGACTGCGTGTGTTACTCCACGTTGGGACCGACGACGAGAACTCCGGCGCTGCA 426
Db 801 CACCTATTTCGTGCGGATCCTCACCGAGGTTCAGCGGCTTCCTCCGCTCGACGCTACCGCT 860

Qy 427 TATCGTTTGGGACTCGTGGCCATGATGTGCTCAGTCCACCAACAGACACT 486
Db 861 GCTGCTGTGTGTACTACGCGCTTGGGACGCTGGTGGGCAACACATCGTCGCGCCCTGGC 920

Qy 487 CTTTCGCTACTCCGTGTCCTCGTGGGCTACGGCTTCTACGGGACATC 535
Db 921 CGACCGCATACCATCGCGGTCTTGCCCTTCGACCTGCTGGCGGCCATC 969

RESULT 8
US-09-621-976-16656
; Sequence 16656, Application US/09621976
; Patent No. 6839063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16656
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16656

Query Match          4.3%; Score 42.2; DB 4; Length 430;
Best Local Similarity 19.2%; Pred. No. 0.014;
Matches 50; Conservative 99; Mismatches 112; Indels 0; Gaps 0;

Qy 640 CTTCCCTGCACACACAGCTGGGATCTCCAGGGATAGGAGCCCTCGCGGCGAGGATG 699
Db 35 MNCMMRSCCYWCMCGRRSCSRAMCCYYKKKSCSRAMCCCYTYCKSCSSYKGS 94

Qy 700 CTTTCTTGGAGGAAAGAACGACGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 759
Db 95 YTTKRAMMWRKRRSCYTSRRRRYYWRSYWMRSMWKGSCCCSCGSGSCYKKKKKGKGC 154
```

```
Qy 760 TTTGGAAGCTCGGAGGACGTGGAGGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCAT 819
Db 155 CMRSYWMCCYYRARRRMWKGSGSCHMYTFRMRMCCCCMRRRSMRRRRCMVKGSYTY 214

Qy 820 CAATGCCACAAACATGCTCTGTGTCGCCGAGAGCCCGAGGGGCTCTCCCGGGTGC 879
Db 215 CYKSSSMCMWARRKRRARGKKRMCCYTKGGGRMMKYCCMRKGRRACTCTGTTCACTGC 274

Qy 880 CCACCTGGGAGACGGGTCTTC 900
Db 275 CGTGTCTGCCAGAAGGCCTTC 295

RESULT 9
US-09-902-540-9500
; Sequence 9500, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 9500
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-9500

Query Match          4.2%; Score 40.8; DB 4; Length 1014;
Best Local Similarity 46.2%; Pred. No. 0.064;
Matches 135; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

Qy 215 GGCATCGTCTGTGTCGGGGAGATGGTATGTTACGCGAGGTGCTGCACGGTCTGATGGG 274
Db 679 GGCAGCGCGCGCTGGGCAACCCGTGGATCTTCGCGAGCTCTCGGAGGCGCCCGCCGC 738

Qy 275 AGGACGACAGGAGCGCGGCTGCACAGAACACACCCCGGCTGTGCTGTCGCCAGT 334
Db 739 AGCCCGCGGAGCGCTCGAGCTGTGCTGGAGACATTCGGGCGCACCTGGACTTCATG 798

Qy 335 AGCCTCCGGATTGGAATCATTCGCCAGGGTCAACGAGCTGCTGTGTTACTCCACCGTG 394
Db 799 GGGGACCGCTGGCGCGCTGCTCTTCGCAAGCAGTGGCGTGTGTACGCCACCGGC 858

Qy 395 GGACACGACGACGAGAAACCTCGCGCTGCATATCGTTGTTGGGGACTCGCTGGCCATG 454
Db 859 TTGTACGCGCGCGCGCTTCCTCGTGGAGGTGAACCGCTTGGACCTGCGCGTCAAGCGGTG 918

Qy 455 GATGTGCTCTCAGTCCACCAACAGCACACTCTCTCGCTACTCCGTGTCCTCC 506
Db 919 GAGGACTGCGTGGCGCGCTTCTTCGCGCGGCCCATGTGGACCTCGCGGGCC 970

RESULT 10
US-09-902-540-1080/c
; Sequence 1080, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
```

Qy	335	AGCCTCGGATTGGAATCATTTCCCGCAGGGTCAA	CGGACTGCGTGTACTCCACCGTG	394
Db	721	GGCTCGGGCAACGGCTTGGCCGAGCAGTGAAC	CAGCA CGGGGATTGAGCCAGCCCTG	780
Qy	395	GGCACAGCGACGAGAAACCTCGGCGTGCATAT	CGTTGTTGGGGAAGTGCCTGGGCATG	454
Db	781	GGCCTCGACCTGTTGTCTCAACTGCTCACTGT	TGTCCTGTGCCGGGTGTGTGCCACCCACTG	840
Qy	455	GATGTGTCCTCAGTCCACCACACAGCACTCT	CTTCGCTACTCCGCTGTCCTGCTGGC	514
Db	841	GACCTGCTCTCCGTGAGCGCTCGGGCTCCG	CTGTTCTCTCTGTCGTGTGGCC	900
Qy	515	TACGGCTTCTACCGGGACATCATCAAGGAC	AGTGAAGAAACGGTGTGGGCTCTTGCC	574
Db	901	TGGGGCTTCGTGTCAAGTGGATATCCAGAG	GAGGGCTTCAGGGCCTTGGGCACTGCC	960
Qy	575	AGATAGACATTTTCAGGTTTTAAGACCTTCT	CTCCCAACATGCTATGAAGGACAGTG	634
Db	961	CGCTTCACACTGGGACGGTGTGGGCTCGC	CACACTGCACCTACCGCGACGCTC	1020
Qy	635	TCCTTCTCCTCCCTGC	648	
Db	1021	TCCTACCTCCCCGC	1034	
RESULT 12				
US-09-817-676A-13				
; Sequence 13, Application US/09817676A				
; Patent No. 6800470				
; GENERAL INFORMATION:				
; APPLICANT: Spiegel, Sarah				
; APPLICANT: Kohama, Takafumi				
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,				
; FILE REFERENCE: Expression and Methods of Use Thereof				
; CURRENT APPLICATION NUMBER: US/09/817,676A				
; CURRENT FILING DATE: 2001-03-26				
; PRIOR APPLICATION NUMBER: US 60/194,318				
; PRIOR FILING DATE: 2000-04-03				
; NUMBER OF SEQ ID NOS: 15				
; SOFTWARE: PatentIn Ver. 2.0				
; SEQ ID NO 13				
; LENGTH: 2380				
; TYPE: DNA				
; ORGANISM: Homo sapiens				
; FEATURE:				
; NAME/KEY: CDS				
; LOCATION: (7)..(1860)				
; PUBLICATION INFORMATION:				
; TITLE: Molecular cloning and functional characterization of a				
; TITLE: novel mammalian sphingosine kinase type 2 isoform				
; JOURNAL: J. Biol. Chem.				
; VOLUME: 275				
; ISSUE: 26				
; PAGES: 19513-19520				
; DATE: 2000-06-30				
; DATABASE ACCESSION NUMBER: AF245447				
; DATABASE ENTRY DATE: 2000-06-27				
US-09-817-676A-13				
Query Match 4.1%; Score 40.4; DB 4; Length 2380;				
Best Local Similarity 44.3%; Pred. No. 0.15;				
Matches 219; Conservative 0; Mismatches 266; Indels 9; Gaps 1;				
Qy	155	ACTGAACATGCTAATCAGGCCAAGGAGACTCT	GTGTATGAGATTAAATAGACAAATACGAC	214
Db	556	ACAGAACGACAGAACACGCCGGAGCTGGTCC	AGGGGCTGAGCTGAGTGAAGTGGAT	615
Qy	215	GGCATCGTCTGTGTCCGCGGAGATGTTATGTT	CAGGAGGTGCTGCACGGTCTGATTGGG	274
Db	616	GGCATCGTCAACGGTCTCGGGAGACGGGCT	GCTCCATGAGGTGCTGAACCGGCTCCTAGAT	675

275	AGGACGAGGAGGAGCGCGGGTGCACAGAACACCCCGGGCTGTGTGTTCCCACTG	334
676	CGCCCTGACTGGGAGGAGCTGTGAAGATG-----CCTGTGGGCATCTCTCCCTGC	726
335	AGCTCCGGATTGGAATCATTTCCCGCAGGGTCAACGGACTGCGTGTGTTACTCCACCGTG	394
727	GGCTCGGGCAACCGCGTGGCCGGAGCAGTGAACGAGCACGGGGATTGTGACCGAGCCCTG	786
395	GGCACCAAGCACGACGAGAAACCTCGGCGCTGCATATCGTTGTGGGGACTCGCTGGCCATG	454
787	GGCCTCGACTGTGTCTCAACTGCTCACTGTTGTGTGCCGGGTGTGTGCCACCCACTG	846
455	GATGTGCTCAGTCCACCAACAGACACTCTCTTCGTACTCCGTGTTCCTGCTGGCC	514
847	GACCTGCTCCGTGAGCTGGCCCTCGGGCTCCGGCTGTTTCTCCTTCCTGTGTGGCC	906
515	TACGGCTTCTACGGGGACATCATCAAGGACAGTACAGAAACCGTGGTTGGGCTTTGCC	574
907	TGGGGCTTCGTGTCCAGATGTGGATATCCAGAGCAGCGCTTTCAGGGCCCTTGGGCAGTGCC	966
575	AGATACGACTTTTCAGGTTTAAAGACTTCTCCTCTCCACCACTGCTATAGAGGGACAGTG	634
967	CGCTTCACACTGGGCACGGTGCTGGGCTCGCCACACTGCACACTACCGCGGAGCGCTC	1026
635	TCCTTCTCCTCGCTGC	648
1027	TCCTACTCTCCCGC	1040

RESULT 13

```

US-09-270-767-31476/c
; Sequence 31476, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31476
; LENGTH: 485
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-31476

```

## RESULT 14

```

US-09/817-676A-11
; Sequence 11, Application US/09817676A
; Patent No. 6800470
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; TITLE OF INVENTION: Expression and Methods of Use Thereof
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817,676A
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/194,318

```

```

; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2698
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (387)..(2237)
; PUBLICATION INFORMATION:
; TITLE: Molecular cloning and functional characterization of a
; TITLE: novel mammalian sphingosine kinase type 2 isoform
; JOURNAL: J. Biol. Chem.
; VOLUME: 275
; ISSUE: 26
; PAGES: 19513-19520
; DATE: 2000-06-30
; DATABASE ACCESSION NUMBER: AF245448
; DATABASE ENTRY DATE: 2000-06-27
; US-09-817-676A-11

```

## RESULT 15

```

US-09-614-221A-399
; Sequence 399, Application US/09614221A
; Patent No. 6723837
; GENERAL INFORMATION:
; APPLICANT: Karunanandaa, Balasulojini
; APPLICANT: Yu, Jaehyuk
; APPLICANT: Kishore, Ganesh M.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED
; TITLE OF INVENTION: WITH STEROL SYNTHESIS AND METABOLISM
; FILE REFERENCE: 16516. 075
; CURRENT APPLICATION NUMBER: US/09/614,221A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/142,981
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 626
; SEQ ID NO 399
; LENGTH: 1875
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
; US-09-614-221A-399

```

Qy 275 AGGACGCAGAGAGCGCCGGGGTCCA 300  
|||  
Db 916 AGACCCGACAGAGTGGATGCGTTCAA 941  
|||

Search completed: September 5, 2005, 15:12:11  
Job time : 189.716 secs



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2005, 12:52:06 ; Search time 711.691 Seconds  
(without alignments)  
9008.663 Million cell updates/sec

Title: US-10-631-958-1

Perfect score: 979

Sequence: 1 accaagcatttactgtat.....acaccaaccagcaggaccag 979

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7338684 seqs, 3274456166 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database : Published Applications NA:\*
- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
  - 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
  - 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
  - 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
  - 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
  - 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
  - 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
  - 8: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
  - 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
  - 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
  - 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
  - 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
  - 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
  - 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
  - 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
  - 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*
  - 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*
  - 18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*
  - 19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq:\*
  - 20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq:\*
  - 21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq:\*
  - 22: /cgn2\_6/ptodata/2/pubpna/US10J\_PUBCOMB.seq:\*
  - 23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq:\*
  - 24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*
  - 25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
  - 26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	979	100.0	979	10	US-09-969-896-1
2	979	100.0	979	19	US-10-631-958-1
3	979	100.0	1840	9	US-09-784-810A-5
4	979	100.0	1840	22	US-10-876-281-5
5	903	92.2	1614	10	US-09-969-896-9
6	903	92.2	1614	19	US-10-631-958-9
7	903	92.2	1740	18	US-10-262-511-39

8	903	92.2	4413	10	US-09-969-896-16	Sequence 16, Appl
9	903	92.2	4413	19	US-10-631-958-16	Sequence 16, Appl
10	903	92.2	4429	19	US-10-618-941-55	Sequence 55, Appl
11	901.4	92.1	4432	17	US-10-120-988-148	Sequence 148, App
12	898.2	91.7	4463	16	US-10-315-597A-1	Sequence 1, Appli
13	464.4	47.4	474	10	US-09-969-896-4	Sequence 4, Appli
14	464.4	47.4	474	19	US-10-631-958-4	Sequence 4, Appli
15	302.2	30.9	329	19	US-09-969-896-5	Sequence 5, Appli
16	302.2	30.9	329	19	US-10-631-958-5	Sequence 5, Appli
17	200.6	20.5	564	22	US-10-477-445-46	Sequence 46, Appl
18	167	17.1	167	10	US-09-969-896-6	Sequence 6, Appli
19	167	17.1	167	19	US-10-631-958-6	Sequence 6, Appli
20	153	15.6	153	10	US-09-969-896-7	Sequence 7, Appli
21	153	15.6	153	19	US-10-631-958-7	Sequence 7, Appli
22	140	14.3	817	19	US-10-115-635-247	Sequence 247, App
c 23	136.8	14.0	522	9	US-09-784-810A-7	Sequence 7, Appli
c 24	136.8	14.0	522	22	US-10-876-281-7	Sequence 7, Appli
25	89	9.1	550	10	US-09-969-896-8	Sequence 8, Appli
26	89	9.1	550	19	US-10-631-958-8	Sequence 8, Appli
c 27	69.8	7.1	498	9	US-09-893-737-59	Sequence 59, Appl
28	65	6.6	1774	18	US-10-425-114-34081	Sequence 34081, A
29	65	6.6	2657	20	US-10-425-115-176413	Sequence 176413,
30	59.6	6.1	861	20	US-10-425-115-86219	Sequence 86219, A
31	54.8	5.6	2079	18	US-10-425-114-34530	Sequence 34530, A
32	54.8	5.6	2079	20	US-10-425-115-86211	Sequence 86211, A
c 33	51	5.2	763	22	US-10-477-445-49	Sequence 49, Appl
c 34	43.2	4.4	1758	20	US-10-425-115-101679	Sequence 101679,
35	43	4.4	1869	18	US-10-425-114-34806	Sequence 34806, A
36	43	4.4	2166	20	US-10-425-115-64890	Sequence 64890, A
37	41.8	4.3	1167	18	US-10-389-647-211	Sequence 211, App
38	41.2	4.2	1161	15	US-10-156-761-7033	Sequence 7033, Ap
39	41.2	4.2	9025608	15	US-10-156-761-1	Sequence 1, Appli
40	40.4	4.1	1857	9	US-09-970-516-3	Sequence 3, Appli
41	40.4	4.1	1857	18	US-10-619-344-3	Sequence 13, Appl
42	40.4	4.1	2380	9	US-09-817-676A-13	Sequence 13, Appl
43	40.4	4.1	2380	16	US-10-354-358-77	Sequence 77, Appl
44	40.4	4.1	2380	19	US-10-283-975A-515	Sequence 515, App
45	40.4	4.1	2380	20	US-10-830-677-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1

US-09-969-896-1

; Sequence 1, Application US/09969896

; Publication No. US20030125533A1

; GENERAL INFORMATION:

; APPLICANT: Kossida, Sophia

; TITLE OF INVENTION: Regulation of human Sphingosine

; FILE REFERENCE: 004974.00594

; CURRENT APPLICATION NUMBER: US/09/969,896

; CURRENT FILING DATE: 2001-10-04

; PRIOR APPLICATION NUMBER: US 60/238,005

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: US 60/314,113

; PRIOR FILING DATE: 2001-08-23

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 979

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-969-896-1

Query Match	100.0%;	Score 979;	DB 10;	Length 979;
Best Local Similarity	100.0%;	Pred. No. 1e-308;		
Matches 979;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ACCAAGCATTACTGCTATTATCAACCCGTTTGGAGGAAAGGACAAAGCGCGAT	60	
Db	1	ACCAAGCATTACTGCTATTATCAACCCGTTTGGAGGAAAGGACAAAGCGCGCGAT	60	

```
Qy 61 ATATGAAGAAAGTGGCAACACGTGTTACCTTAGCTCCATCACCACATGACATCATCGG 120
Db 61 ATATGAAGAAAGTGGCAACACGTGTTACCTTAGCTCCATCACCACATGACATCATCGG 120
Qy 121 TAACAAATTCATGTTAACTATGTAGAGTAATTAAGTGAACATGCTAATCAGGCAAGGA 180
Db 121 TAACAAATTCATGTTAACTATGTAGAGTAATTAAGTGAACATGCTAATCAGGCAAGGA 180
Qy 181 GACTCTGTATGAGATTAAACATAGACAAATACGACGGGATCGTGTGCGCGAGATGG 240
Db 181 GACTCTGTATGAGATTAAACATAGACAAATACGACGGGATCGTGTGCGCGAGATGG 240
Qy 241 TATGTTACGAGGCTGCTGCAACGTCTGATTGGGAGGACGAGAGGCGCGGGGTGCA 300
Db 241 TATGTTACGAGGCTGCTGCAACGTCTGATTGGGAGGACGAGAGGCGCGGGGTGCA 300
Qy 301 CCAGAACCAACCCCGGGCTGTGCTGCCAGTAGCCTCCGGATTGGAATCATTCGCGC 360
Db 301 CCAGAACCAACCCCGGGCTGTGCTGCCAGTAGCCTCCGGATTGGAATCATTCGCGC 360
Qy 361 AGGTCAACGGAATGCTGCTGTTACTCCACCGTGGGCAACGACGACGAGAACTCGGC 420
Db 361 AGGTCAACGGAATGCTGCTGTTACTCCACCGTGGGCAACGACGACGAGAACTCGGC 420
Qy 421 GCTGCATATCGTGTGGGGACTCGCTGGCCATGGATGTCCTCAGTCCACCAACACAG 480
Db 421 GCTGCATATCGTGTGGGGACTCGCTGGCCATGGATGTCCTCAGTCCACCAACACAG 480
Qy 481 CACACTCTTCGCTACTCCGTTCCCTGCTGGGCTACGGCTTCTACGGGACATCAAA 540
Db 481 CACACTCTTCGCTACTCCGTTCCCTGCTGGGCTACGGCTTCTACGGGACATCAAA 540
Qy 541 CGACGTGAGAGAAACGGTGGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAGAC 600
Db 541 CGACGTGAGAGAAACGGTGGTGGTCTTCCAGATACGACTTTTCAGGTTTAAAGAC 600
Qy 601 CTTCTCTCCCAACACTGCTATGAAGGACAGTGTCTTCTCCTCGTGCACAAACACAGGT 660
Db 601 CTTCTCTCCCAACACTGCTATGAAGGACAGTGTCTTCTCCTCGTGCACAAACACAGGT 660
Qy 661 GGGATCTCCAAAGGATAGAAAGCCCTGCGGCGAGGATGCTTTGTTGAGGCAAAAGCAA 720
Db 661 GGGATCTCCAAAGGATAGAAAGCCCTGCGGCGAGGATGCTTTGTTGAGGCAAAAGCAA 720
Qy 721 GCAGCAGCTGAGAGGACGAGAAAGCACTGTATGGTTTGAAGCTGCGGAGGACGT 780
Db 721 GCAGCAGCTGAGAGGACGAGAAAGCACTGTATGGTTTGAAGCTGCGGAGGACGT 780
Qy 781 GGAGAGTGGCAAGTCTGTGGGAAGTTTCTGGCCATCAATGCGCAAAACATGTCTCTG 840
Db 781 GGAGAGTGGCAAGTCTGTGGGAAGTTTCTGGCCATCAATGCGCAAAACATGTCTCTG 840
Qy 841 TGCTTGTCCCGAGCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTC 900
Db 841 TGCTTGTCCCGAGCCCCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTC 900
Qy 901 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTCTCATCAGGCA 960
Db 901 TGACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTCTCATCAGGCA 960
Qy 961 CACCAACACGACGAGGACACAG 979
Db 961 CACCAACACGAGGACACAG 979
```

RESULT 2

```
US-10-631-958-1
; Sequence 1, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
```

```
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 979
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-958-1
```

```
Query Match 100.0%; Score 979; DB 19; Length 979;
Best Local Similarity 100.0%; Pred. No. 1e-308;
Matches 979; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCAAGCATTTACTGTTATTTATCAACCCGTTTGGAGGAAAGGACAAGGCAAGCGGAT 60
Db 1 ACCAAGCATTTACTGTTATTTATCAACCCGTTTGGAGGAAAGGACAAGGCAAGCGGAT 60
Qy 61 ATATGAAGAAAGTGGCAACACGTGTTACCTTAGCTCCATCACCACATGACATCATCGG 120
Db 61 ATATGAAGAAAGTGGCAACACGTGTTACCTTAGCTCCATCACCACATGACATCATCGG 120
Qy 121 TAACAAATTCATGTTAACTATGTAGAAAGTAATTAAGTGAACATGCTAATCAGGCAAGGA 180
Db 121 TAACAAATTCATGTTAACTATGTAGAAAGTAATTAAGTGAACATGCTAATCAGGCAAGGA 180
Qy 181 GACTCTGTATGAGATTAAACATAGACAAATACGACGGGATCGTGTGTCGGCGAGATGG 240
Db 181 GACTCTGTATGAGATTAAACATAGACAAATACGACGGGATCGTGTGTCGGCGAGATGG 240
Qy 241 TATGTTACGAGGCTGCTGCAACGTCTGATTGGGAGGACGAGAGGACGCGGGGTGCA 300
Db 241 TATGTTACGAGGCTGCTGCAACGTCTGATTGGGAGGACGAGAGGACGCGGGGTGCA 300
Qy 301 CCAGAACCAACCCCGGGCTGTGCTGGTCCCAGTAGCTCCGGATTGGAATCATTCGCGC 360
Db 301 CCAGAACCAACCCCGGGCTGTGCTGGTCCCAGTAGCTCCGGATTGGAATCATTCGCGC 360
Qy 361 AGGTCAACGGAATGCTGCTGTTACTCCACCGTGGGCAACGACGACGAGAACTCGGC 420
Db 361 AGGTCAACGGAATGCTGCTGTTACTCCACCGTGGGCAACGACGACGAGAACTCGGC 420
Qy 421 GCTGCATATCGTGTGGGGACTCGCTGGCCATGGATGTCCTCAGTCCACCAACACAG 480
Db 421 GCTGCATATCGTGTGGGGACTCGCTGGCCATGGATGTCCTCAGTCCACCAACACAG 480
Qy 481 CACACTCTTCGCTACTCCGTTCCCTGCTGGGCTACGGCTTCTACGGGACATCAAA 540
Db 481 CACACTCTTCGCTACTCCGTTCCCTGCTGGGCTACGGCTTCTACGGGACATCAAA 540
Qy 541 GGACAGTGAAGAAACGGTGGTGGTCTTGGCCATCAATGCGCAAAACATGTCTCTG 600
Db 541 GGACAGTGAAGAAACGGTGGTGGTCTTGGCCATCAATGCGCAAAACATGTCTCTG 600
Qy 601 CTTCTCTCCCAACACTGCTATGAAGGACAGTGTCTTCTCCTCGTGCACAAACACAGGT 660
Db 601 CTTCTCTCCCAACACTGCTATGAAGGACAGTGTCTTCTCCTCGTGCACAAACACAGGT 660
Qy 661 GGGATCTCCAAAGGATAGAAAGCCCTGCGGCGAGGATGCTTTGTTGAGGCAAAAGCAA 720
Db 661 GGGATCTCCAAAGGATAGAAAGCCCTGCGGCGAGGATGCTTTGTTGAGGCAAAAGCAA 720
Qy 721 GCAGCAGCTGAGAGGACGAGAAAGCACTGTATGGTTTGAAGCTGCGGAGGACGT 780
Db 721 GCAGCAGCTGAGAGGACGAGAAAGCACTGTATGGTTTGAAGCTGCGGAGGACGT 780
```

QY	781	GGAGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCCTG	840
Db	781	GGAGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCCTG	840
QY	841	TGCTTGTGCGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTC	900
Db	841	TGCTTGTGCGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTC	900
QY	901	TGACCTCATCTCATCCCGGAATGCTCCAGGTTCAATTTCTGAGATTCTCATCAGGCA	960
Db	901	TGACCTCATCTCATCCCGGAATGCTCCAGGTTCAATTTCTGAGATTCTCATCAGGCA	960
QY	961	CACCAACCAGCAGGACCAG	979
Db	961	CACCAACCAGCAGGACCAG	979



```
Db 1134 GGAGGAGTGGCAAGTCTGTGGGAAGTTTCTGGCCATCAATGCCAACAACATGCTCTG 1193
Qy 841 TGCTTGTGCGCGAGGCCAGGGGCTCTCCCGGCTGCCACTTTGGGAGCGGTCTTC 900
Db 1194 TGCTTGTGCGCGAGGCCAGGGGCTCTCCCGGCTGCCACTTTGGGAGCGGTCTTC 1253
Qy 901 TGACCTCATCTTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTCTCATCAGGA 960
Db 1254 TGACCTCATCTTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTCTCATCAGGA 1313
Qy 961 CACCAACAGCAGGACCAG 979
Db 1314 CACCAACAGCAGGACCAG 1332

RESULT 6
US-10-631-958-9
; Sequence 9, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; PRIOR FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1614
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-958-9
```

```
Query Match 92.2%; Score 903; DB 19; Length 1614;
Best Local Similarity 96.6%; Pred. No. 9.3e-284;
Matches 946; Conservative 0; Mismatches 0; Indels 33; Gaps 1;

Qy 1 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGGAAAGCAAGCAAGCGGAT 60
Db 387 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGGAAAGCAAGCAAGCGGAT 446
Qy 61 ATATGAAAGAAAGTGGCAACCACTGTTTACCTTAGCCTCCATCACCACCTGACATCGG 120
Db 447 ATATGAAAGAAAGTGGCAACCACTGTTTACCTTAGCCTCCATCACCACCTGACATCGG 505
Qy 121 TAACAAATTCATGTTAACTATGATAGAGTAATTAAGTCAATGCTAATCAGGCAAGGA 180
Db 506 -----TTACTGAACATGCTAATCAGGCAAGGA 533
Qy 181 GACTCTGATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGCGCGAGATGG 240
Db 534 GACTCTGATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGCGCGAGATGG 593
Qy 241 TATGTTACGAGAGTGTGCGACGGTCTGATTTGGGAGGACGACAGAGAGCGCGGGGTGCA 300
Db 594 TATGTTACGAGAGTGTGCGACGGTCTGATTTGGGAGGACGACAGAGAGCGCGGGGTGCA 653
Qy 301 CCAGAACCAACCCCGGCTGTGCTGCTCCAGTAGTCCCGATTGGAATCATTCCTCCGC 360
Db 654 CCAGAACCAACCCCGGCTGTGCTGCTCCAGTAGTCCCGATTGGAATCATTCCTCCGC 713
Qy 361 AGGGTCAACGGACTCGTGTGTTACTTCCACCGTGGGCAACGACGACGAGAAACCTCGGC 420
Db 714 AGGGTCAACGGACTCGTGTGTTACTTCCACCGTGGGCAACGACGACGAGAAACCTCGGC 773
```

```
Qy 421 GCTCATATCGTTGTTGGGGACTCGTGGCCATGATGTGTCTCAGTCCACCAACAG 480
Db 774 GCTCATATCGTTGTTGGGGACTCGTGGCCATGATGTGTCTCAGTCCACCAACAG 833
Qy 481 CACACTCTTCGCTACTCCGTGTCCTGCTGGGCTTACGGCTTCTACGGGACATCATCAA 540
Db 834 CACACTCTTCGCTACTCCGTGTCCTGCTGGGCTTACGGCTTCTACGGGACATCATCAA 893
Qy 541 GGACAGTGAAGAAACGGTGGTTCCTTGGCAGATACGACTTTTTCAGGTTTAAAGAC 600
Db 894 GGACAGTGAAGAAACGGTGGTTCCTTGGCAGATACGACTTTTTCAGGTTTAAAGAC 953
Qy 601 CTTCTCTCCCAACCACTGCTATGAAGGACAGTGTCTTCTCCTCTGACACACACCGT 660
Db 954 CTTCTCTCCCAACCACTGCTATGAAGGACAGTGTCTTCTCCTCTGACACACACCGT 1013
Qy 661 GGGATCTCAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTTCAGGCAAGCAA 720
Db 1014 GGGATCTCAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTTCAGGCAAGCAA 1073
Qy 721 GCAGCAGCTGGAGGAGGAGCAAGAAAGCACTGTATGTTTGGAAAGCTGCGGAGCGT 780
Db 1074 GCAGCAGCTGGAGGAGGAGCAAGAAAGCACTGTATGTTTGGAAAGCTGCGGAGCGT 1133
Qy 781 GGAGAGTGGCAAGTCTGTGTGGAAAGTTTCTGCCCATCAATGCCAACAACATGTCTG 840
Db 1134 GGAGGAGTGGCAAGTCTGTGTGGAAAGTTTCTGCCCATCAATGCCAACAACATGTCTG 1193
Qy 841 TGCTTGTGCGCGAGGAGCCCGAGGGGCTCTCCCGGCTGCCCCACATTTGGGAGCGGTCTTC 900
Db 1194 TGCTTGTGCGCGAGGAGCCCGAGGGGCTCTCCCGGCTGCCCCACATTTGGGAGCGGTCTTC 1253
Qy 901 TGACCTCATCTCATCCGGAATGCTCCAGGTTTCAATTTTCTGAGATTCTCATCAGGCA 960
Db 1254 TGACCTCATCTCATCCGGAATGCTCCAGGTTTCAATTTTCTGAGATTCTCATCAGGCA 1313
Qy 961 CACCAACAGCAGGACCAG 979
Db 1314 CACCAACAGCAGGACCAG 1332

RESULT 7
US-10-262-511-39
; Sequence 39, Application US/10262511
; Publication No. US20040038223A1
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glenna
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Rameeh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Saaha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly R.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zernusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
```

APPLICANT: Agee, Michele L.  
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: 21402-462C  
CURRENT APPLICATION NUMBER: US/10/262,511  
PRIORITY FILING DATE: 2003-05-28  
PRIORITY FILING DATE: 2001-10-02  
PRIORITY FILING DATE: 2001-10-02  
PRIORITY FILING DATE: 2002-04-19  
PRIORITY FILING DATE: 2002-04-19  
PRIORITY FILING DATE: 2001-10-09  
PRIORITY FILING DATE: 2002-05-17  
PRIORITY FILING DATE: 2002-05-17  
PRIORITY FILING DATE: 2002-10-09  
PRIORITY FILING DATE: 2002-05-16  
PRIORITY FILING DATE: 2002-04-19  
PRIORITY FILING DATE: 2001-10-09  
PRIORITY FILING DATE: 2001-10-05  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 439  
SOFTWARE: CuraseqList version 0.1  
SEQ ID NO 39  
LENGTH: 1740  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (76)..(1686)  
US-10-262-511-39

Query Match 92.2%; Score 903; DB 18; Length 1740;  
Best Local Similarity 96.6%; Pred. No. 9.6e-284; Indels 33; Gaps 1;  
Matches 946; Conservative 0; Mismatches 0

Qy 1 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGAAAAGGACAAAGCGGAT 60  
Db 462 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGAAAAGGACAAAGCGGAT 521

Qy 61 ATATGAAAGAAAGTGGCACCACCTGTTCACTTACCTCCATCAACCACTGACATCATCGG 120  
Db 522 ATATGAAAGAAAGTGGCACCACCTGTTCACTTACCTCCATCAACCACTGACATCATCG- 580

Qy 121 TAACAAATTCATGTTAACTATGTAGAAGTAATTACTGAACATGCTAATCAGGCCAAGGA 180  
Db 581 -----TTACTGAACATGCTAATCAGGCCAAGGA 608

Qy 181 GACTCTGTATGAGATTAAACATAGACAAATACGCGCATCGTGTGTCGCGGAGATGG 240  
Db 609 GACTCTGTATGAGATTAAACATAGACAAATACGCGCATCGTGTGTCGCGGAGATGG 668

Qy 241 TATGTTACGAGGTGTCGACCGTCTGATTGGAGGACGACAGAGAGCGCCGGGTGCA 300  
Db 669 TATGTTACGAGGTGTCGACCGTCTGATTGGAGGACGACAGAGAGCGCCGGGTGCA 728

Qy 301 CCAGAACACCCCGGGTGTGCTGCTCCAGTAGCTCCGAGTTGGAATCATTTCCCGC 360  
Db 729 CCAGAACACCCCGGGTGTGCTGCTCCAGTAGCTCCGAGTTGGAATCATTTCCCGC 788

Qy 361 AGGGTCAACGAGCTCGTGTGTTACTCCACCGTGGGACACGACGACGAGAACTCCGC 420  
Db 789 AGGGTCAACGAGCTCGTGTGTTACTCCACCGTGGGACACGACGACGAGAACTCCGC 848

Qy 421 GCTGATATCGTTGTTGGGACTCGCTGGCCATGATGTGCTCAGTCCACCAACAG 480  
Db 849 GCTGATATCGTTGTTGGGACTCGCTGGCCATGATGTGCTCAGTCCACCAACAG 908

Qy 481 CACACTCTTCGCTACTCCGTCCTGCTGGGTACGGCTTCTACGGGACATCATCAA 540  
Db 909 CACACTCTTCGCTACTCCGTCCTGCTGGGTACGGCTTCTACGGGACATCATCAA 968

Qy 541 GGCAGTGAGAAAGAAACGGTGGTCTTGCAGATACGACTTTTTCAGGTTTAAAGAC 600  
Db 969 GGCAGTGAGAAAGAAACGGTGGTCTTGCAGATACGACTTTTTCAGGTTTAAAGAC 1028

Qy 601 CTTCTCTCCACCACTGCTATGAAAGGACAGTGTCTTCTCCCTGCAACACACGGT 660  
Db 1029 CTTCTCTCCACCACTGCTATGAAAGGACAGTGTCTTCTCCCTGCAACACACGGT 1088

Qy 661 GGGATCTCCAAGGATAGGAAGCCCTCGCGGGCAGGATGCTTTGTTGAGGCAAAACAA 720  
Db 1089 GGGATCTCCAAGGATAGGAAGCCCTCGCGGGCAGGATGCTTTGTTGAGGCAAAACAA 1148

Qy 721 GCAGCAGCTCGAGGAGGAGCAGAAAGCACTGTATGTTTGGAACTCGCGGAGGACGT 780  
Db 1149 GCAGCAGCTCGAGGAGGAGCAGAAAGCACTGTATGTTTGGAACTCGCGGAGGACGT 1208

Qy 781 GAGGAGTGCAGAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCTG 840  
Db 1209 GAGGAGTGCAGAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCTG 1268

Qy 841 TGCTTGTGCGCGGAGCCCGCAGGGCCCTCTCCCGGCTGCCACTTTGGGAGACGGGTCTTC 900  
Db 1269 TGCTTGTGCGCGGAGCCCGCAGGGCCCTCTCCCGGCTGCCACTTTGGGAGACGGGTCTTC 1328

Qy 901 TGACCTCATCTCATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 960  
Db 1329 TGACCTCATCTCATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 1388

Qy 961 CACCAACACGACGACCCAG 979  
Db 1389 CACCAACACGACGACCCAG 1407

RESULT 8  
US-09-969-896-16  
; Sequence 16, Application US/09969896  
; Publication No. US20030125533A1  
; GENERAL INFORMATION:  
; APPLICANT: Kossida, Sophia  
; TITLE OF INVENTION: Regulation of human Sphingosine  
; FILE REFERENCE: 004974.00594  
; CURRENT APPLICATION NUMBER: US/09/969,896  
; CURRENT FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: US 60/238,005  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: US 60/314,113  
; PRIOR FILING DATE: 2001-08-23  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 4413  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-969-896-16

Query Match 92.2%; Score 903; DB 10; Length 4413;  
Best Local Similarity 96.6%; Pred. No. 1.6e-283;  
Matches 946; Conservative 0; Mismatches 0; Indels 33; Gaps 1;  
Qy 1 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGAAAAGGACAAAGCGGAT 60  
Db 462 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGAAAAGGACAAAGCGGAT 521

Qy 61 ATATGAAAGAAAGTGGCACCACCTGTTCACTTACCTCCATCACCACCTGACATCATCGG 120  
Db 522 ATATGAAAGAAAGTGGCACCACCTGTTCACTTACCTCCATCACCACCTGACATCATCG- 580



Qy	121	TAA	CAAA	TCTTATGTTAACTATGTAGAGAGTAATTA	TTACTGAA	CAATGCTAATCAGGCGCAAGGA	180
Db	581	-----	-----	-----	TTACTGAA	CATGCTAATCAGGCGCAAGGA	608
Qy	181	GACTCTGTATCAGATTAACATAGACAAATACGACGGCATCGTCTGTGTGCGCGGAGATGG	240				
Db	609	GACTCTGTATCAGATTAACATAGACAAATACGACGGCATCGTCTGTGTGCGCGGAGATGG	668				
Qy	241	TATGTTGACGAGGTGCTGCACGGTCTGATTTGGGAGGACGACGAGGAGCGCGGGGTGCA	300				
Db	669	TATGTTGACGAGGTGCTGCACGGTCTGATTTGGGAGGACGACGAGGAGCGCGGGGTGCA	728				
Qy	301	CCAGAACACCCCGGGCTGTGCTGCTCCACAGTAGCCTCCGGATTGGAAATCATTTCCCGC	360				
Db	729	CCAGAACACCCCGGGCTGTGCTGCTCCACAGTAGCCTCCGGATTGGAAATCATTTCCCGC	788				
Qy	361	AGGGTCAACGCACTCGCTGTGTTACTCCACCGTGGGCACCAAGCGACGAGAAACCTCGGC	420				
Db	789	AGGGTCAACGCACTCGCTGTGTTACTCCACCGTGGGCACCAAGCGACGAGAAACCTCGGC	848				
Qy	421	GCTGCATATCGTTGTTGGGACTCGCTGGGCCATGGATGTGTCTCAGTCCACCAACAG	480				
Db	849	GCTGCATATCGTTGTTGGGACTCGCTGGGCCATGGATGTGTCTCAGTCCACCAACAG	908				
Qy	481	CACACTCCTTCGCTACTCCGTGTCCCTCTGGGCTACCGGCTTCTACGGGGACATCATCAA	540				
Db	909	CACACTCCTTCGCTACTCCGTGTCCCTCTGGGCTACCGGCTTCTACGGGGACATCATCAA	968				
Qy	541	GGACAGTAGAAGAAAACGGTGGTTGGGTCTTGGCAGATACGACTTTTCAGGTTTAAAGAC	600				
Db	969	GGACAGTAGAAGAAAACGGTGGTTGGGTCTTGGCAGATACGACTTTTCAGGTTTAAAGAC	1028				
Qy	601	CTTCCTCTCCCAACCACTGCTATGAAGGAGACAGTGTCTTCTCCTCGCACAAACACAGGT	660				
Db	1029	CTTCCTCTCCCAACCACTGCTATGAAGGAGACAGTGTCTTCTCCTCGCACAAACACAGGT	1088				
Qy	661	GGGATCTCCAAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTTGCAGGCGCAAGCAA	720				
Db	1089	GGGATCTCCAAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTTGCAGGCGCAAGCAA	1148				
Qy	721	GCACGACTGGAGGAGGACGAGAAGCACTGTATGTTTGGNAGCTGCGGAGGAGCT	780				
Db	1149	GCACGACTGGAGGAGGACGAGAAGCACTGTATGTTTGGNAGCTGCGGAGGAGCT	1208				
Qy	781	GGAGGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGGCCATCAATGCCACAAATGTCCTG	840				
Db	1209	GGAGGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGGCCATCAATGCCACAAATGTCCTG	1268				
Qy	841	TGCTTGTGCGCGGAGCCCGAGGGGCTCTCCCGGCTGCCCATTTGGGAGACGGGTCTTC	900				
Db	1269	TGCTTGTGCGCGGAGCCCGAGGGGCTCTCCCGGCTGCCCATTTGGGAGACGGGTCTTC	1328				
Qy	901	TGACCTCATCTCATCCGAAATGCTCCAGGTTCAAATTTCTGAGATTTCTCATCAGGCA	960				
Db	1329	TGACCTCATCTCATCCGAAATGCTCCAGGTTCAAATTTCTGAGATTTCTCATCAGGCA	1388				
Qy	961	CACCAACGACGAGCACG	979				
Db	1389	CACCAACGACGAGCACG	1407				

## RESULT 9

```

US-10-631-958-16
:
: Sequence 16, Application US/10631958
: Publication No. US20040192580A1
:
: GENERAL INFORMATION:
:
: APPLICANT: Kossida, Sophia
: TITLE OF INVENTION: Regulation of human S
: TITLE OF INVENTION: Kinase-Like Protein
: FILE REFERENCE: 004974_00594
: CURRENT APPLICATION NUMBER: US/10/631,958
: CURRENT FILING DATE: 2003-08-01
: PRIOR APPLICATION NUMBER: US/09/969,896
:

```

Qy 841 TGCTTGTCCGCGAGCCCCAGGGGCTCTCCCGCTGCCCCATCTGGGAGACGGGTCTTC 900  
Db 1269 TGCTTGTCCGCGAGCCCCAGGGGCTCTCCCGCTGCCCCATCTGGGAGACGGGTCTTC 1328  
Qy 901 TGACCTCATCTCATTCGGAAATGCTCAGGTTCAATTTTCTGAGTTTCTCATCAGGCA 960  
Db 1329 TGACCTCATCTCATTCGGAAATGCTCAGGTTCAATTTTCTGAGTTTCTCATCAGGCA 1388  
Qy 961 CACCAACAGCAGGACCG 979  
Db 1389 CACCAACAGCAGGACCG 1407  
RESULT 10  
US-10-618-941-55  
; Sequence 55, Application US/10618941  
; Publication No. US20040197792A1  
; GENERAL INFORMATION:  
; APPLICANT: MANNING, GERARD  
; APPLICANT: CAENEPEEL, SEAN  
; TITLE OF INVENTION: NOVEL KINASES  
; FILE REFERENCE: 034536-0321  
; CURRENT APPLICATION NUMBER: US/10/618,941  
; CURRENT FILING DATE: 2003-07-15  
; PRIOR FILING DATE: 60/395,632  
; PRIOR FILING DATE: 2002-07-15  
; NUMBER OF SEQ ID NOS: 143  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 55  
; LENGTH: 4429  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-618-941-55

Query Match 92.2%; Score 903; DB 19; Length 4429;  
Best Local Similarity 96.6%; Pred. No. 1.6e-283;  
Matches 946; Conservative 0; Mismatches 0; Indels 33; Gaps 1;  
Qy 1 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGGAAAGACAGGCAAGCGGAT 60  
Db 478 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGGAAAGACAGGCAAGCGGAT 537  
Qy 61 ATATGAAAGAAAGTGGCACCACCTGTTCACTTAGCTCCATCACCACCTGACATCATCGG 120  
Db 538 ATATGAAAGAAAGTGGCACCACCTGTTCACTTAGCTCCATCACCACCTGACATCATCG- 596  
Qy 121 TAACAAATTTCTATGTTAACTATGTAGAAAGTAAATCTGAAACATGCTAATCAGGCCAAGGA 180  
Db 597 -----TTACTGAACATGCTAATCAGGCCAAGGA 624  
Qy 181 GACTCTGTATGATTAACATAGACAAATACGACGGCATCTCTGTGTCGGGGAGATGG 240  
Db 625 GACTCTGTATGATTAACATAGACAAATACGACGGCATCTCTGTGTCGGGGAGATGG 684  
Qy 241 TATGTTACAGCAGGTGCTGCAGGTCTGATTTGGGAGGACGACAGAGCGCGCGGGTCTGA 300  
Db 685 TATGTTACAGCAGGTGCTGCAGGTCTGATTTGGGAGGACGACAGAGCGCGCGGGTCTGA 744  
Qy 301 CCAGAACCAACCCCGGGCTGTGCTGGTCCCAGTAGCTCCCGGATTTGGAATCATTTCCCGC 360  
Db 745 CCAGAACCAACCCCGGGCTGTGCTGGTCCCAGTAGCTCCCGGATTTGGAATCATTTCCCGC 804  
Qy 361 AGGTCACAGGACTCGGTGTTTACTCCACCGTGGGACCGACGACGAGAAACCTCGGC 420  
Db 805 AGGTCACAGGACTCGGTGTTTACTCCACCGTGGGACCGACGACGAGAAACCTCGGC 864  
Qy 421 GCTGCATATCGTTGTTGGGACTCGCTGGCCATGATGTCTCCTCAGTCCACCAACAG 480  
Db 865 GCTGCATATCGTTGTTGGGACTCGCTGGCCATGATGTCTCCTCAGTCCACCAACAG 924  
Qy 481 CACACTCTCTGCTACTCTCGGTGCTCGGTGAGGCTTACCGGCTTCTACGGGACATCATCA 540

Db 925 CACACTCTCTCGCTACTCCGTTGTCCTGCTGGGCTACGGCTTCTTACGGGACATCATCAA 984  
Qy 541 GGACAGTGAAGAAACGGTGGGTCTTGGCAGATACGACTTTTTCAGGTTTAAAGAC 600  
Db 985 GGACAGTGAAGAAACGGTGGGTCTTGGCAGATACGACTTTTTCAGGTTTAAAGAC 1044  
Qy 601 CTTCTCTCTCCACCATCTGCTATGAAGGACAGTGTCTCTCTCCCTGCACAACACACGGT 660  
Db 1045 CTTCTCTCTCCACCATCTGCTATGAAGGACAGTGTCTCTCTCCCTGCACAACACACGGT 1104  
Qy 661 GGGATCTCAAGGATAGGAAGCCCTCCGGGACAGGATGCTTTGTTGCGAGGCAAGCA 720  
Db 1105 GGGATCTCAAGGATAGGAAGCCCTCCGGGACAGGATGCTTTGTTGCGAGGCAAGCA 1164  
Qy 721 GCAGCAGCTGGAGGAGCAGAGAAAGCACTGTATGGTTTGAAGCTCGGAGGACGT 780  
Db 1165 GCAGCAGCTGGAGGAGCAGAGAAAGCACTGTATGGTTTGAAGCTCGGAGGACGT 1224  
Qy 781 GGAGGAGTGGCAAGTCTCTGTGGGAAGTTTCTGGCCATCAATGSCCAACAAATGTCTG 840  
Db 1225 GGAGGAGTGGCAAGTCTCTGTGGGAAGTTTCTGGCCATCAATGSCCAACAAATGTCTG 1284  
Qy 841 TGCTTGTGCGCGAGAGCCCCAGGGGCTCTCTCCCGGCTGCGCACTTGGGAGACGGGTCTTC 900  
Db 1285 TGCTTGTGCGCGAGAGCCCCAGGGGCTCTCTCCCGGCTGCGCACTTGGGAGACGGGTCTTC 1344  
Qy 901 TGACTCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 960  
Db 1345 TGACTCTCATCTCATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCA 1404  
Qy 961 CACCAACAGCAGGACCG 979  
Db 1405 CACCAACAGCAGGACCG 1423

RESULT 11  
US-10-988-148  
; Sequence 148, Application US/10120988  
; Publication No. US20030219745A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Goodrich, Ryle  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. US20030219745A1el Nucleic Acids and  
; FILE REFERENCE: Polypeptides  
; FILE REFERENCE: 802CON  
; CURRENT APPLICATION NUMBER: US/10/120,988  
; CURRENT FILING DATE: 2002-04-11  
; PRIOR FILING DATE: 09/774,528  
; PRIOR FILING DATE: 2001-01-30  
; NUMBER OF SEQ ID NOS: 441  
; SOFTWARE: pt\_FL\_genes Version 2.0  
; SEQ ID NO 148  
; LENGTH: 4432  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1129) .. (2817)  
US-10-988-148

Query Match 92.1%; Score 901.4; DB 17; Length 4432;  
Best Local Similarity 96.5%; Pred. No. 5.3e-283;  
Matches 945; Conservative 0; Mismatches 1; Indels 33; Gaps 1;  
Qy 1 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGGAAAGCAAGCAAGCGGAT 60  
Db 1590 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGGAAAGCAAGCAAGCGGAT 1649  
Qy 61 ATATGAAAGAAAGTGGCAACCATGTTTCACTTAGCCTTCATCACCCTGATCATCGG 120

```

|||||
Db 1650 ATATGAAAGAAAGTGGACCACTGTTCACCTTAGCCTCCATCCACCTGACATCATCG- 1708
Qy 121 TAACAAATTTCTATGTAACTATGTAGAAGTAATTAAGTGAACATGTAATTAAGGCAAGGA 180
Db 1709 -----TTACTGAACATGCTAATCAGGCCAAGGA 1736
Qy 181 GACTCTGTATGAGATTAAATAGACAAATAAGACGGCATCGTCTGTGTGCGGGAGATGG 240
Db 1737 GACTCTGTATGAGATTAAATAGACAAATAAGACGGCATCGTCTGTGTGCGGGAGATGG 1796
Qy 241 TATGTTTCAAGGAGTGTGTCAGGTCATTTGGAGGACGACGAGAGCGCGGGGTGCA 300
Db 1797 TATGTTTCAAGGAGTGTGTCAGGTCATTTGGAGGACGACGAGAGCGCGGGGTGCA 1856
Qy 301 CCAGAACCAACCCCGGGGTGTCTGGTCCCAAGTAGTCCGATTCGGATTTGGAATCATTTCCGCG 360
Db 1857 CCAGAACCAACCCCGGGGTGTCTGGTCCCAAGTAGTCCGATTCGGATTTGGAATCATTTCCGCG 1916
Qy 361 AGGTCACAGGACTCGGTGTGTTACTTCAACCGTGGGACACGACGAGCAAACTCGGC 420
Db 1917 AGGTCACAGGACTCGGTGTGTTACTTCAACCGTGGGACACGACGAGCAAACTCGGC 1976
Qy 421 GCTGCATATCGTTGTTGGGACTCGCTGSCCATGTATGTCTCAGTCCACCAACAG 480
Db 1977 GCTGCATATCGTTGTTGGGACTCGCTGSCCATGTATGTCTCAGTCCACCAACAG 2036
Qy 481 CACACTCTCTGCTACTCCGTGTCTGCTGGGCTACGGCTTCTACGGGGACATCATCA 540
Db 2037 CACACTCTCTGCTACTCCGTGTCTGCTGGGCTACGGCTTCTACGGGGACATCATCA 2096
Qy 541 GGACAGTGAGAGAAACGGTGTGGTCTTGGCTTGGCCAGATACGACTTTTTCAGGTTTAAAGAC 600
Db 2097 GGACAGTGAGAGAAACGGTGTGGTCTTGGCTTGGCCAGATACGACTTTTTCAGGTTTAAAGAC 2156
Qy 601 CTTCTCTCCACCACTCTATGAGGACAGTGTCTTCTCTCCCTGACACACACAGGT 660
Db 2157 CTTCTCTCCACCACTCTATGAGGACAGTGTCTTCTCTCCCTGACACACACAGGT 2216
Qy 661 GGGATCTCCAAAGGATAGAAAGCCCTCGCGGACAGTGTCTTGTTCAGGCAAGCA 720
Db 2217 GGGATCTCCAAAGGATAGAAAGCCCTCGCGGACAGTGTCTTGTTCAGGCAAGCA 2276
Qy 721 GCAGAGCTGGAGAGAGAGAGAGAGAAAGCACTGTATGTTTGGAGCTGCGGAGAGCT 780
Db 2277 GCAGAGCTGGAGAGAGAGAGAGAGAAAGCACTGTATGTTTGGAGCTGCGGAGAGCT 2336
Qy 781 GGAGAGTGGCAAGTCTGTGGAAGTCTTGGCCATCAATGCCACAAACATGTCCTG 840
Db 2337 GGAGAGTGGCAAGTCTGTGGAAGTCTTGGCCATCAATGCCACAAACATGTCCTG 2396
Qy 841 TGCTTGTGCGGAGCGCCAGGGGCTCTCCCGGCTGCCCCACTTTGGGAGACGGGTCTTC 900
Db 2397 TGCTTGTGCGGAGCGCCAGGGGCTCTCCCGGCTGCCCCACTTTGGGAGACGGGTCTTC 2456
Qy 901 TGACCTCATCTCATCTCGGAAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGCA 960
Db 2457 TGACCTCATCTCATCTCGGAAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGCA 2516
Qy 961 CACCAACACGAGGACCCAG 979
Db 2517 CACCAACACGAGGACCCAG 2535
```

RESULT 12

US-10-315-597A-1

; Sequence 1, Application US/10315597A  
; Publication No. US20030162206A1  
; GENERAL INFORMATION:  
; APPLICANT: Sugiyura, Masako  
; APPLICANT: Kono, Keita  
; APPLICANT: Kohama, Takafumi  
; TITLE OF INVENTION: Ceramide Kinase and DNA Encoding It

```

; FILE REFERENCE: 02658CIP/HG
; CURRENT APPLICATION NUMBER: US/10/315.597A
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: JP 2000-178039
; PRIOR FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4463
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 3371
; OTHER INFORMATION:
; US-10-315-597A-1

Query Match 91.7%; Score 898.2; DB 16; Length 4463;
Best Local Similarity 96.3%; Pred. No. 5.9e-282;
Matches 943; Conservative 0; Mismatches 3; Indels 33; Gaps 1;

Qy 1 ACCAAAGCATTTACTGTGTTATTTATCAACCCGTTTGGAGGAAAAGGACAAAGCGGAT 60
Db 510 ACCAAAGCATTTACTGTGTTATTTATCAACCCGTTTGGAGGAAAAGGACAAAGCGGAT 569
Qy 61 ATATGAAAGAAAAGTGGCACCACTGTTCACTTAGCCTTCCATCACCCTGACATCATCGG 120
Db 570 ATATGAAAGAAAAGTGGCACCACTGTTCACTTAGCCTTCCATCACCCTGACATCATCG- 628
Qy 121 TAACAAATTCATGTAACTATGTAGAAGTAATTTACTGAACATGCTAATCAGGCCAAGGA 180
Db 629 -----TTACTGAACATGCTAATCAGGCCAAGGA 656
Qy 181 GACTCTGTATGAGATTAAATAGACAAATAGACAGGCGCATCGTCTGTGTCGGCGAGATGG 240
Db 657 GACTCTGTATGAGATTAAATAGACAAATAGACAGGCGCATCGTCTGTGTCGGCGAGATGG 716
Qy 241 TATGTTACGAGAGTGTGCAACGGTCTGATTGGAGAGACGACAGAGAGCGCGGGGTGCA 300
Db 717 TATGTTACGAGAGTGTGCAACGGTCTGATTGGAGAGACGACAGAGAGCGCGGGGTGCA 776
Qy 301 CCAGAACCAACCCCGGGCTGTGTCCTCCAGTAGCCTCCGGATTGGAATCATTTCCCGC 360
Db 777 CCAGAACCAACCCCGGGCTGTGTCCTCCAGTAGCCTCCGGATTGGAATCATTTCCCGC 836
Qy 361 AGGTCACAGGACTCGGTGTGTTACTCCACCGTGGGACACGACGACGAGAAACCTCGGC 420
Db 837 AGTCCAAAGGACTCGGTGTGTTACTCCACCGTGGGACACGACGACGAGAAACCTCGGC 896
Qy 421 GCTGCATATCGTTGTTGGGACTCGTGGCCATGGATGTGTCCTCAGTCCACCAACAG 480
Db 897 GCTGCATATCGTTGTTGGGACTCGTGGCCATGGATGTGTCCTCAGTCCACCAACAG 956
Qy 481 CACACTCTTCGCTACTCCGTGTCCTGCTGGGCTACGGCTTCTACGGGACATCATCA 540
Db 957 CACACTCTTCGCTACTCCGTGTCCTGCTGGGCTACGGCTTCTACGGGACATCATCA 1016
Qy 541 GGACAGTGAGAGAAAACGGTGTGGGTCTTTGCGAGATACGACTTTTTCAGGTTTAAAGAC 600
Db 1017 GGACAGTGAGAGAAAACGGTGTGGGTCTTTGCGAGATACGACTTTTTCAGGTTTAAAGAC 1076
Qy 601 CTTCTCTTCCCAACCATGTCTATGAAGGACAGTGTCTTCTCCTGTCACCAACACAGGT 660
Db 1077 CTTCTCTTCCCAACCATGTCTATGAAGGACAGTGTCTTCTCCTGTCACCAACACAGGT 1136
Qy 661 GGGATCTCCAGGATAGGAGCCCTCGCGGACAGGATGCTTTTTCAGGCAAGCA 720
Db 1137 GGGATCTCCAGGATAGGAGCCCTCGCGGACAGGATGCTTTTTCAGGCAAGCA 1196
Qy 721 GCAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGAAGCTGCGGAGAGCT 780
Db 1197 GCAGCAGCTGGAGGAGGACAGAAAGCACTGTATGTTTGAAGCTGCGGAGAGCT 1256
```

```
Qy 781 GGAGGATGGCAAGTCGTCTCTGGGAAGTTTCTGGCCATCAATGCCAACAACATGTCTTG 840
Db 1257 GGAGGATGGCAAGTCGTCTCTGGGAAGTTTCTGGCCATCAATGCCAACAACATGTCTTG 1316

Qy 841 TGCTTTGTCGCGGAGCCCCAGGGGCTTCTCCCGGCTGCCCCACTTGGGAGACGGGTCTTC 900
Db 1317 TGCTTTGTCGCGGAGCCCCAGGGGCTTCTCCCGGCTGCCCCACTTGGGAGACGGGTCTTC 1376

Qy 901 TGACCTCATCTCTATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTCTCATCAGGCA 960
Db 1377 TGACCTCATCTCTATCCGGAATGCTCCAGGTTCAATTTTCTGAGATTCTCATCAGGCA 1436

Qy 961 CACCAACAGCAGGACCAG 979
Db 1437 CACCAACAGCAGGACCAG 1455

RESULT 13
US-09-969-896-4
; Sequence 4, Application US/09969896
; Publication No. US20030125533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-4
```

```
Query Match 47.4%; Score 464.4; DB 10; Length 474;
Best Local Similarity 99.8%; Pred. No. 1.2e-140;
Matches 465; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 239 GGTATGTTTCAGCGAGTCTGCACGGTCTGATTGGGAGGACGACAGAGCGCGGGGTC 298
Db 8 GGTATGTTTCAGCGAGTCTGCACGGTCTGATTGGGAGGACGACAGAGCGCGGGGTC 67

Qy 299 GACCAGAACCAACCCCGGGCTGTGTGTCGCCAGTAGCCTCCGGATTGGAATCATTTCCC 358
Db 68 GACCAGAACCAACCCCGGGCTGTGTGTCGCCAGTAGCCTCCGGATTGGAATCATTTCCC 127

Qy 359 GCAGGGTCAACGAGTCTGGTGTGTACTCCACCGTGGGACACGAGCGAGAAACCTCG 418
Db 128 GCAGGGTCAACGAGTCTGGTGTGTACTCCACCGTGGGACACGAGCGAGAAACCTCG 187

Qy 419 GCGTGCATATCGTTGTTGGGACTCGCTGCCATGGATGTCTCAGTCCACCAAC 478
Db 188 GCGTGCATATCGTTGTTGGGACTCGCTGCCATGGATGTCTCAGTCCACCAAC 247

Qy 479 AGCACACTCTTCGCTACTCCGTGTCCCTGTGGGCTACGGCTTCTACGGGACATCATC 538
Db 248 AGCACACTCTTCGCTACTCCGTGTCCCTGTGGGCTACGGCTTCTACGGGACATCATC 307

Qy 539 AAGGACAGTAGAAGAAACCGGTGGTTCGCCAGATACGACTTTTCAGGTTTAAAG 598
Db 308 AAGGACAGTAGAAGAAACCGGTGGTTCGCCAGATACGACTTTTCAGGTTTAAAG 367

Qy 599 ACCTTCCTCTCCACCACTGTATGAAGGACAGTGTCTTCTCCCTGACACACACG 658
Db 368 ACCTTCCTCTCCACCACTGTATGAAGGACAGTGTCTTCTCCCTGACACACACG 427

Qy 659 GTGGGATCTCCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTG 704
```

```
Db 428 GTGGGATCTCCAAGGATAGGAAGCCCTGCGGCAAGATGCTTTG 473

RESULT 14
US-10-631-958-4
; Sequence 4, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-958-4
```

```
Query Match 47.4%; Score 464.4; DB 19; Length 474;
Best Local Similarity 99.8%; Pred. No. 1.2e-140;
Matches 465; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 239 GGTATGTTTCAGCGAGTCTGCACGGTCTGATTGGGAGGACGACAGAGCGCGGGGTC 298
Db 8 GGTATGTTTCAGCGAGTCTGCACGGTCTGATTGGGAGGACGACAGAGCGCGGGGTC 67

Qy 299 GACCAGAACCAACCCCGGGCTGTGTGTCGCCAGTAGCCTCCGGATTGGAATCATTTCCC 358
Db 68 GACCAGAACCAACCCCGGGCTGTGTGTCGCCAGTAGCCTCCGGATTGGAATCATTTCCC 127

Qy 359 GCAGGGTCAACGAGTCTGGTGTGTACTCCACCGTGGGACACGAGCGAGAAACCTCG 418
Db 128 GCAGGGTCAACGAGTCTGGTGTGTACTCCACCGTGGGACACGAGCGAGAAACCTCG 187

Qy 419 GCGTGCATATCGTTGTTGGGACTCGCTGCCATGGATGTCTCAGTCCACCAAC 478
Db 188 GCGTGCATATCGTTGTTGGGACTCGCTGCCATGGATGTCTCAGTCCACCAAC 247

Qy 479 AGCACACTCTTCGCTACTCCGTGTCCCTGTGGGCTACGGCTTCTACGGGACATCATC 538
Db 248 AGCACACTCTTCGCTACTCCGTGTCCCTGTGGGCTACGGCTTCTACGGGACATCATC 307

Qy 539 AAGGACAGTAGAAGAAACCGGTGGTTCGCCAGATACGACTTTTCAGGTTTAAAG 598
Db 308 AAGGACAGTAGAAGAAACCGGTGGTTCGCCAGATACGACTTTTCAGGTTTAAAG 367

Qy 599 ACCTTCCTCTCCACCACTGTATGAAGGACAGTGTCTTCTCCCTGACACACACG 658
Db 368 ACCTTCCTCTCCACCACTGTATGAAGGACAGTGTCTTCTCCCTGACACACACG 427

Qy 659 GTGGGATCTCCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTG 704
Db 428 GTGGGATCTCCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTG 473
```

```
RESULT 15
US-09-969-896-5
; Sequence 5, Application US/09969896
; Publication No. US20030125533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
```

```
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 329
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(329)
; OTHER INFORMATION: n = A,T,C or G
US-09-969-896-5
```

```
Query Match      30.9%; Score 302.2; DB 10; Length 329;
Best Local Similarity 97.7%; Pred. No. 1e-87;
Matches 304; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 153 TTACTGAACATGCTTAATCAGGCCAAGGAGACTCTCTATGAGATTAAACATAGACAAATACG 212
Db 19 TTACTGAACATGCTTAATCAGGCCAAGGAGACTCTCTATGAGATTAAACATAGACAAATACG 78

QY 213 ACGGCATCGTCTGTCTCGCGGAGATGGTATGTTTCAGCGAGGTCTGCACCGTCTGATTG 272
Db 79 ACGGCATCGTCTGTCTCGCGGAGATGGTATGTTTCAGCGAGGTCTGCACCGTCTGATTG 138

QY 273 GGAGACGCGAGAGCGCGGGGTGACCCAGAACACCCCGGGCTGTGTGGTCCCA 332
Db 139 GGAGACGCGAGAGCGCGGGGTGACCCAGAACACCCCGGGCTGTGTGGTCCCA 198

QY 333 GTAGCCTCGGATTGGAATCATTCGCCAGGTCACCGACTGCGTGTGTTACTCCACG 392
Db 199 GTAGCCTCGGATTGGAATCATTCGCCAGGTCACCGACTGCGTGTGTTACTCCACG 258

QY 393 TGGGCACCGACGACGAGAAACCTCGGCGTGCATATGTTGTTGGGACTCGCTGGCCA 452
Db 259 TGGGCACCGACGACGAGAAACCTCGGCGTGCATATGTTGTTGGGACTCGCTGGCCA 318

QY 453 TGGATGTGTCC 463
Db 319 TGGATGTGTCC 329
```

Search completed: September 5, 2005, 20:00:08  
Job time : 726.691 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2005, 18:18:34 ; Search time 3579.6 Seconds  
(without alignments)  
10410.363 Million cell updates/sec

Title: US-10-631-958-1

Perfect score: 979  
Sequence: 1 accaaagattactgtat.....acacaaaccagcaggaccag 979

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hc1:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gse1:\*  
9: gb\_gse2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	801.4	81.9	1063	BM479389	BM479389 AGENCOURT
2	726.4	74.2	1059	BQ054406	BQ054406 AGENCOURT
3	696.6	71.2	797	CK000755	CK000755 AGENCOURT
4	670	68.4	1047	BQ057191	BQ057191 AGENCOURT
5	635.8	64.9	1824	AK042077	AK042077 Mus muscu
6	635.8	64.9	4248	AK052269	AK052269 Mus muscu
7	593.4	60.6	820	CD655311	CD655311 AGENCOURT
8	582.4	59.5	1078	BQ063738	BQ063738 AGENCOURT
9	579.4	59.2	581	BP224560	BP224560 BP224560
10	537.4	54.9	758	CB246749	CB246749 UI-HF-BNO
11	536.2	54.8	584	BP310011	BP310011 BP310011
12	535.8	54.7	541	CN296311	CN296311 17005321
13	527.4	53.9	584	BP309990	BP309990 BP309990
14	520.8	53.2	573	CF138275	CF138275 UI-HF-BNO
15	518.2	52.9	550	CF138634	CF138634 UI-HF-BNO
16	505	51.6	713	CK000525	CK000525 AGENCOURT
17	503	51.4	661	CN296312	CN296312 17005321
18	477.2	48.7	488	BW430459	BW430459 UI-HF-BNO
19	464.4	47.4	474	AW503999	AW503999 UI-HF-BNO
20	445	45.5	653	BX952302	BX952302 DKFZp781L
21	415	42.4	732	CF135528	CF135528 UI-HF-BNO
22	402.8	41.1	576	CB608292	CB608292 AMGNNUC:N
23	367.8	37.6	545	CB611947	CB611947 AMGNNUC:N
24	367.4	37.5	376	BU430866	BU430866 UI-HF-BNO

25	348.8	35.6	359	5	BU428926	BU428926 UI-HF-BNO
26	340	34.7	553	6	CA578972	CA578972 K0731A08-
27	318.8	32.6	725	5	BU214294	BU214294 603756937
28	306.2	31.3	674	1	AJ739034	AJ739034 AJ739034
29	302.2	30.9	329	1	AA355581	AA355581 EST64087
30	294	30.0	779	5	BU265377	BU265377 60374309
31	288.6	29.5	467	4	BM484519	BM484519 538413 MA
32	278.8	28.5	763	7	CK364327	CK364327 AGENCOURT
33	278.6	28.5	385	5	BU428707	BU428707 UI-HF-BNO
34	274.8	28.1	902	5	BU373554	BU373554 603811294
35	268.2	27.4	607	6	CB581157	CB581157 AMGNNUC:N
36	267.2	27.3	452	6	CA579952	CA579952 K0744G05-
37	263.2	26.9	735	7	CF535441	CF535441 UI-M-GH0-
38	260	26.6	310	6	CD631568	CD631568 56059637H
39	258.4	26.4	310	6	CD631566	CD631566 56059545H
40	253.6	25.9	759	5	BU290807	BU290807 604162346
41	242	24.7	291	6	CD631569	CD631569 56059637J
42	239.8	24.5	292	6	CD631567	CD631567 56059545J
43	234.4	23.9	759	5	BU264162	BU264162 603814452
44	222	22.7	567	1	AJ739023	AJ739023 AJ739023
45	215.6	22.0	638	1	AJ739022	AJ739022 AJ739022

ALIGNMENTS

RESULT 1  
BM479389  
LOCUS BM479389 1063 bp mRNA linear EST 05-FEB-2002  
DEFINITION AGENCOURT 6418742 NIH\_MGC\_67 Homo sapiens cDNA clone IMAGE:5502281  
5', mRNA sequence.  
ACCESSION BM479389  
VERSION BM479389.1 GI:18528431  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1063)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM12140 row: c column: 18  
High quality sequence stop: 665.  
FEATURES  
Location/Qualifiers  
1..1063  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5502281"  
/tissue\_type="retinoblastoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_67"  
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.75 kb. Library constructed by Life Technologies."

ORIGIN  
Query Match 81.9%; Score 801.4; DB 4; Length 1063;  
Best Local Similarity 98.7%; Pred. No. 5.2e-219;  
Matches 805; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
QY 153 TTACTGAACATGCTAATCAGCCAGGAGACTCTGTATGAGATTACATAGACAATACG 212





QY 609 CCCACACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACAAACACACGCGTGGATCTC 668  
|||||  
Db 580 CCCACACTGCTATGAAGGACAGTGTCTTCTCTCCCTGCACAAACACACGCGTGGATCTC 639  
QY 669 CAAGGATAGGAAGCCCTCCCGGAGGATGCTTTGTTGAGGCAAGACAGCAGCAGC 728  
Db 640 CAAGGATAGGAAGCCCTCCCGGAGGATGCTTTGTTGAGGCAAGACAGCAGCAGC 699  
QY 729 TGGAGGAGGAGCAGAAAGACACTGTATGTTTGGAAAGCTCGGAGGACCTGGAGGAGT 788  
Db 700 TGGAGGAGGAGCAGAAAGACACTGTATGTTTGGAAAGCTCGGAGGACCTGGAGGAGT 759  
QY 789 GCGAAGTCTCTGTGGG-AAGTTTCTGGCCATCATGCCCACAA 831  
Db 760 GCGAAGTCTCTGTGGGAAGTTCTTGGCCATTAATGCCACAA 803

RESULT 3  
CK000755  
LOCUS  
DEFINITION AGENCOURT\_16363467 NIH\_MGC\_221 Homo sapiens cDNA clone  
IMAGE:30707875 5', mRNA sequence.  
ACCESSION CK000755  
VERSION CK000755.1 GI:38526789  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.  
1 (bases 1 to 797)  
NIH-MGC <http://mgs.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: [gcapbs-remail.nih.gov](mailto:gcapbs-remail.nih.gov)  
Tissue Procurement: James Martin, University of Iowa  
cDNA Library Preparation: M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: NDAM1073 row: C column: 20  
High quality sequence stop: 656.  
Location/Qualifiers  
1..797  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30707875"  
/lab\_host="DH10B Tona"  
/clone\_lib="NIH\_MGC\_221"  
/note="Organ: mixed; Vector: pYX-Asc; Site 1: EcoRI;  
Site 2: NotI; Library is oligo-dT primed and directionally  
cloned. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT  
primer containing a Not I site. Double strand cDNA was  
size selected according to mRNA size fraction, ligated with  
EcoR I adaptor, digested with Not I and then cloned  
directionally into pYX-Asc vector. Average insert size  
4-5Kb. Adaptors 5'(AATTGGCAGCAGG)3' and 5'd  
(CTGTGTCGG)3'. 3' Linker sequence - CGGCGCTGAGACC T18.  
Sequencing primers 3'end: T3 promoter primer 5'd  
(ATTAACTCTCAATAAGGA)3'. 5' End: T7 promoter primer 5'd  
(TAATACGACTCATATAGG)3'. Library was constructed in the  
laboratory of M. Bento Soares. Note: this is a NIH\_MGC  
Library"

FEATURES  
source

QY 813 TGGCCATCAATGCCACAAACATGCTGCTTGTGCGCGAGGCCCTCTCC 872  
Db 721 TGGCCNATCATGCACAAACATGCTGCTTGTGCGCGAGGCCCTCTCCCGG 780  
QY 873 CGCG 876  
Db 781 CTGC 784

## ORIGIN

RESULT 4  
LOCUS BQ057191  
DEFINITION AGENCOURT\_6769595 NIH\_MGC\_99 Homo sapiens cDNA clone IMAGE:581282  
5', mRNA sequence.  
ACCESSION BQ057191  
VERSION BQ057191.1 GI:19816531  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.  
1 (bases 1 to 1047)  
NIH-MGC <http://mgs.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)

Query Match 71.2%; Score 696.6; DB 7; Length 797;  
Best Local Similarity 97.4%; Pred. No. 7e-189;  
Matches 705; Conservative 0; Mismatches 19; Indels 0; Gaps 0;  
QY 153 TTACTGAACATGCTTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACG 212  
Db 61 TTACTGAACATGCTTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACG 120  
QY 213 ACGCATCGTCTGTGTCGGCGAGATGTTTTCAGCAGAGTCTGTCACCGTCTGATTG 272  
Db 121 ACGCATCGTCTGTGTCGGCGAGATGTTTTCAGCAGAGTCTGTCACCGTCTGATTG 180  
QY 273 GGAGGACGACAGAGCGCGGGTTCACACAGAACCCCGGGCTGTGTCGTCCTCCCA 332  
Db 181 GGAGGACGACAGAGCGCGGGTTCACACAGAACCCCGGGCTGTGTCGTCCTCCCA 240  
QY 333 GTAGCTCCGATTGGAAATCATTCGCGAGGCTCAACGAGCTGGGTGTGTTACTTCCACCG 392  
Db 241 GTAGCTCCGATTGGAAATCATTCGCGAGGCTCAACGAGCTGGGTGTGTTACTTCCACCG 300  
QY 393 TGGCACACGACGACGAGAAACCTCGGCGCTGCATATGTTTGTGGGACTCGCTGGCCA 452  
Db 301 TGGCACACGACGACGAGAAACCTCGGCGCTGCATATGTTTGTGGGACTCGCTGGCCA 360  
QY 453 TGGATGTCTCTAGTCCACACACAGCAGTCTCTCGCTACTCCGCTGCTGCTGG 512  
Db 361 TGGATGTCTCTAGTCCACACACAGCAGTCTCTCGCTACTCCGCTGCTGCTGG 420  
QY 513 GCTACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAAACGTTGGTCTTTG 572  
Db 421 GCTACGGCTTCTACGGGACATCATCAAGGACAGTGAAGAAACGTTGGTCTTTG 480  
QY 573 CCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACACTGCTATGAAGGACAG 632  
Db 481 CCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACACTGCTATGAAGGACAG 540  
QY 633 TGTCTCTCTCTCCGACACACACAGCGTGGATCTCCAGGGATAGGAGCCCTGCCGG 692  
Db 541 TGTCTCTCTCTCCGACACACACAGCGTGGATCTCCAGGGATAGGAGCCCTGCCGG 600  
QY 693 CAGATGCTTTGTTTTCAGGCAAGCAAGCAGCAGTGGAGGAGGAGCAGAGAAAGCAC 752  
Db 601 CAGATGCTTTGTTTTCAGGCAAGCAAGCAGCAGTGGAGGAGGAGCAGAGAAAGCAC 660  
QY 753 TGTATGTTTGAAGCTCCGAGGACGCTGGAGGAGTGGCAAGTCTGTGCGGAAGTTTC 812  
Db 661 TGTATGTTTGAAGCTCCGAGGACGCTGGAGGAGTGGNCAGTCTGTCTGTGGGAAGTTTC 720  
QY 813 TGGCCATCAATGCCACAAACATGCTGCTTGTGCGCGAGGCCCTCTCC 872  
Db 721 TGGCCNATCATGCACAAACATGCTGCTTGTGCGCGAGGCCCTCTCCCGG 780  
QY 873 CGCG 876  
Db 781 CTGC 784

## RESULT 4

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov	
	Tissue Procurement: Lou Staudt cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LICM2062 row: 1 column: 15 High quality sequence stop: 535. Location/Qualifiers 1. 1047	
FEATURES	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5812382" /tissue_type="lymphoma, cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH_MGC_99" /note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."	
	source	
ORIGIN	Query Match 68.4%; Score 670; DB 5: Length 1047; Best Local Similarity 91.3%; Pred. No. 3.5e-181; Matches 825; Conservative 0; Mismatches 35; Indels 44; Gaps 9;	
	9 ATTTACTGGTATTTATCAACCGTGTGGAGAAAGGACGAAGCGGATATATGAAA 68 11 ATTTACTGGTATTTATCAACCGTGTGGAGAAAGGACGAAGCGGATATGAAA 70 69 GAAAGTGGCACCAGTCTACCTTAGCTTCATCCATCACCAGTACATCATCGTAAACAAT 128 71 GAAAGTGGCACCAGTCTACCTTAGCTTCATCCATCACCAGTACATCATCGT----- 121 129 TCTATGTTAACTATGTAGAAAGTAATTACTGAACATGCTAACTCAGCCAAAGGAGACTCTGT 188 122 -----TTACTGAACATGCTAATCAGCCCAAGAGACTCTGT 157 189 ATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGGTATGTCA 248 158 ATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGGAGATGGTATGTCA 217 249 GCGAGGTGCTCGACGGTCTGATTGGGAGGACGACGAGGAGCGCGGGTGCACAGAAC 308 218 GCGAGGTGCTCGACGGTCTGATTGGGAGGACGACGAGGAGCGCGGGTGCACAGAAC 277 309 ACCCCGGGCTGTGCTGTGTCGCCAGTAGCTCCGATTGGAATCATTCGCCAGGGTCAA 368 278 ACCCCGGGCTGTGCTGTGTCGCCAGTAGCTCCGATTGGAATCATTCGCCAGGGTCAA 337 369 CGGACTCGTGTGTTACTCCACCGTGGGCACGACGACGCGAGAAACCTTCGGCGTGCATA 428 338 CGGACTCGTGTGTTACTCCACCGTGGGCACGACGACGCGAGAAACCTTCGGCGTGCATA 397 429 TCGTGTGTTGGGACTCGCTGGCCATGGATGTGCTCAGTCCACCAACAGCACACTCC 488 398 TCGTGTGTTGGGACTCGCTGGCCATGGATGTGCTCAGTCCACCAACAGCACACTCC 457 489 TTCGCTACTCCGTGTCCTGCTGGGCTACGGCTCTACGGGACATCATCAAGACAGTG 548 458 TTCGCTACTCCGTGTCCTGCTGGGCTACGGCTCTACGGGACATCATCAAGACAGTG 517 549 AGAAGAAACGGTGTGGGTCTTCCAGATACGACTTTTCAGGTTTAAAGACCTTCCTCT 608 	
Db	518 AGAAGAAACGGTGTGGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGACCTTCCTCT 577 609 CCACCACTCTCTATGAA-GGGACAGTGTCTCTCTCCCTGCACAAACACACGGTGGGATCT 667 578 CCACCACTCTCTATGAAAGGGGACAGTGTCTCTCTCCCTGCACAAACACACGGTGGGATCT 637 668 -CCAAGGATAGGAAGCCCTG-CCGGGACAGATGCTTTG-TTTCAGGCAAAAGCAAGCAG 724 638 CCAAGGATAGGAAGCCCTG-CCGGGACAGATGCTTTG-TTTCAGGCAAAAGCAAGCAG 697 725 CAGCTGGAGGAGGACGAGAAAGC--ACTGTATGTTTGGAGCTGCGGAGACGTGG 782 698 CAGCTGGAGGAGGACGAGAAAGCACCATTGTATGTTTGGAACTTCGCGAAGACGGGC 757 783 AGGAGTGGCAAGTCGTCT-CTGGGAGTTCCTGCCCATCAATGCCACAAACATGCTCT-G 840 758 AGGAGTGGCAAAATCTTCTGGGGGAGTTCCTGCCCATCATGCCCAAAACCTGTCTCTGG 817 841 TGCTTGTGCGCGGAGCCCGACAGGGGCTCT---CCCGGCTGCCCACCTTGGGAGACGGGTC 897 818 TGCTTGTGCGCGGAGCCCGACAGGGGCTCTTCCCGGCTGCCCAATTTGGGAGAGGG 877 898 TTCT 901 878 TCCT 881	
	RESULT 5 AK042077 LOCUS DEFINITION AK042077.1 GI:26334912 Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched library, clone:A630056D11 product:DA59H18.2 (NOVEL PROTEIN SIMILAR TO HUMAN, MOUSE, YEAST, WORM AND PLANT (PREDICTED) PROTEINS) (FRAGMENT) homolog [Homo sapiens], full insert sequence. AK042077 AK042077.1 GI:26334912 Mus musculus (house mouse) HTC; CAP trapper. Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 92279253 10349636 2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000) 20499374 11042159 3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kiseuani,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 20530913 11076861 4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)	
Qy	ACCESSION VERSION KEYWORDS SOURCE ORGANISM REFERENCE AUTHORS TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE JOURNAL	
	AK042077 AK042077.1 GI:26334912 Mus musculus (house mouse) Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 92279253 10349636 2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000) 20499374 11042159 3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kiseuani,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 20530913 11076861 4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)	









```
Db 431 GAAGTTTCTGCCATCAATGCAAAACATGTCCTGTGTTGTTCGCCGAGCCCAAGGG 490
Qy 865 CCTCTCCCGGCTGCCACTTTGGGAGACGGGTCTTCTGACCTCATCTCTCATCCGGAATG 924
Db 491 CCTCTCCCGGCTGCCACTTTGGGAGACGGGTCTTCTGACCTCATCTCTCATCCGGAATG 550
Qy 925 CTCAGGTTCAATTTTCTGAGATTCTTCATCAGGCACACCAACGAGGACACG 979
Db 551 CTCAGGTTCAATTTTCTGAGATTCTTCATCAGGCACACCAACGAGGACACG 605

RESULT 8
BQ063738
LOCUS BQ063738
DEFINITION AGENCOURT_6873251 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5925382
5', mRNA sequence.
ACCESSION BQ063738
VERSION BQ063738.1 GI:19891754
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1078)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2096 row: p column: 23
High quality sequence stop: 640.
FEATURES
Location/Qualifiers
1..1078
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5925382"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
ORIGIN
Query Match 59.5%; Score 582.4; DB 5; Length 1078;
Best Local Similarity 92.6%; Pred. No. 5.8e-156;
Matches 659; Conservative 0; Mismatches 16; Indels 37; Gaps 3;
Qy 1 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGAAAGACAAAGCGAGCGAT 60
Db 120 ACCAAGCATTTACTGGTATTATCAACCCGTTTGGAGAAAGACAAAGCGAGCGAT 179
Qy 61 ATATGAAGAAAGTGGCACCAGTCTTTCACCTTAGCTCCATCCACTGACATCATCGG 120
Db 180 ATATGAAGAAAGTGGCACCAGTCTTTCACCTTAGCTCCATCCACTGACATCATCG- 238
Qy 121 TAACAAAATTCATGTTAACTATGATAGAAAGTAAATCTAGCAATGCTAATCAGGCAAGGA 180
Db 239 -----TTACTGAACATGCTAATCAGGCAAGGA 266
```

```
Qy 181 GACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGAGATGG 240
Db 267 GACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCGTCTGTGTCGGCGAGATGG 326
Qy 241 TATGTTTCAGGAGGTGCTGCACGGTCTGTATGGGAGGACGACAGAGAGCCGCGGGGTGCGA 300
Db 327 TATGTTTCAGGAGGTGCTGCACGGTCTGTATGGGAGGACGACAGAGAGCCGCGGGGTGCGA 386
Qy 301 CCAGAACCCCGGGGCTGTGTCGTCCTCCAGTACGCTCCGATTTGGAAATCATTCCTCCGC 360
Db 387 CCAGAACCCCGGGGCTGTGTCGTCCTCCAGTACGCTCCGATTTGGAAATCATTCCTCCGC 446
Qy 361 AGGCTCAACCGGACTGCGTGTGTACTCCACCGTGGGACGACGACGACGACAAACCTCCGC 420
Db 447 AGGCTCAACCGGACTGCGTGTGTACTCCACCGTGGGACGACGACGACGACAAACCTCCGC 506
Qy 421 GCTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCTCAGTCCACCAACAG 480
Db 507 GCTGCATATCGTTTGGGACTCGCTGGCCATGGATGTCTCTCAGTCCACCAACAG 566
Qy 481 CACACTCTTCGCTACTCCGCTGCTCCCTGCTGGGCTTACGCGGACATCATCAA 540
Db 567 CACACTCTTCGCTACTCCGCTGCTCCCTGCTGGGCTTACGCGGACATCATCAA 626
Qy 541 GGACAGTGAAGAAGAACGGTGGTGGTCTTGGCATGATACGACTTTTTCAGGTTTAAAGAC 600
Db 627 GGACAGTGAAGAAGAACGGTGGTGGTCTTGGCATGATACGACTTTTTCAGGTTTAAAGAC 686
Qy 601 CTTCTCTCCACACATGCTATGAAGGACAGTGTCTTCTCTCCCTGTCACACACAGG- 659
Db 687 CTTCTCTCCACACATGCTATGAAGGACAGGTCCTCTCTCTCTGTCACACACAGG 746
Qy 660 -TGGGATCTCCAGGGA--TAGGAAGCCCTGCCGGCAGGATGCTTTGTTTG 708
Db 747 TGGGGATCTCCAGGGAATAAGGAAGCCCTGCCGGCAGGATGCTTTTG 798

RESULT 9
BQ224560
LOCUS BQ224560
DEFINITION BQ224560 Sugano cDNA library, lymphocyte Daudi Homo sapiens cDNA
clone DAT01067, mRNA sequence.
ACCESSION BQ224560
VERSION BQ224560.1 GI:52097465
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 581)
AUTHORS Suzuki,Y., Yamashita,R., Hirota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
JOURNAL Contact: Yutaka Suzuki
COMMENT Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yezuki@ims.u-tokyo.ac.jp.
FEATURES
Location/Qualifiers
1..581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DAT01067"
/cell_type="lymphocyte"
/cell_line="Daudi"
/clone_lib="Sugano cDNA library, lymphocyte Daudi"
/note="Burkitt's lymphoma"
ORIGIN
```



Query Match	59.2%;	Score 579.4;	DB 5;	Length 581;
Best Local Similarity	99.8%;	Pred. No. 3.5e-155;		
Matches 580;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0
QY	195	TTAA	CATAGACA	CAAAATACACGCGCATCTGCTGTGTGTCGGCGGAGATGGTATGTTTCAGCGAGG 254
DB	1	TTAA	CATAGACA	CAAAATACGACGGCATCTGCTGTGTGTCGGCGGAGATGGTATGTTTCAGCGAGG 60
QY	255	TGCT	GCACGGTCT	GATTGGGAGGACGCAGAGAGAGCGCCGGGGTTCGACCAAGACCAACCCCC 314
DB	61	TGCT	GCACGGTCT	GATTGGGAGGACGCAGAGAGAGCGCCGGGGTTCGACCAAGACCAACCCCC 120
QY	315	GGG	CTGTGCTGGT	CCCACTAGCTCCGGATGGATCATTTCCCGCAGGGTCAACGGACT 374
DB	121	GGG	CTGTGCTGGT	CCCACTAGCTCCGGATGGATCATTTCCCGCAGGGTCAACGGACT 180
QY	375	CGG	TGTGTTACT	CCACCGTGGGACACGACGACGAGAAACCTCGCGCTGCATATCGTTG 434
DB	181	CGG	TGTGTTACT	CCACCGTGGGACACGACGACGAGAAACCTCGCGCTGCATATCGTTG 240
QY	435	TTG	GGGACTCGCT	GGCCATGGATGTCTCTCAGTCCACCAACAGACACACTCTCTTCGCT 494
DB	241	TTG	GGGACTCGCT	GGCCATGGATGTCTCTCAGTCCACCAACAGACACACTCTCTTCGCT 300
QY	495	ACT	CCGTGTCCCT	GTGGGCTACGGCTTCTAGGGGACATCATCAAGGACAGTGAGAGA 554
DB	301	ACT	CCGTGTCCCT	GTGGGCTACGGCTTCTAGGGGACATCATCAAGGACAGTGAGAGA 360
QY	555	AAC	GGTGTGGG	CTTTCGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCAC 614
DB	361	AAC	GGTGTGGG	CTTTCGCCAGATACGACTTTTTCAGGTTTAAAGACCTTCTCTCCAC 420
QY	615	ACT	GCTATGAAG	GACAGTGTCTTCTCTCCCTGCACACACACACGCTGGGATCTCCAGGG 674
DB	421	ACT	GCTATGAAG	GACAGTGTCTTCTCTCCCTGCACACACACACGCTGGGATCTCCAGGG 480
QY	675	ATAG	AAGCCCTG	CGGGACGATGCTTTGTTTGAGGCAAGCAAGCAGCAGCTGGAGG 734
DB	481	ATAG	AAGCCCTG	CGGGACGATGCTTTGTTTGAGGCAAGCAAGCAGCAGCTGGAGG 540
QY	735	AGG	AGCAGAGAA	AGCACTGATGGTTTGGAGCTGCGGAG 775
DB	541	AGG	AGCAGAGAA	AGCACTGATGGTTTGGAGCTGCGGAG 581
RESULT 10				
CB246749				
LOCUS	CB246749	758 bp	mRNA	linear EST 09-JUL-2003
DEFINITION	UI-M-F10-cdx-b-10-0-UI.r1 NIH BMAP_F10 Mus musculus cDNA clone			
IMAGE	IMAGE:6835595 5', mRNA sequence.			
ACCESSION	CB246749			
VERSION	CB246749.1 GI:28368393			
KEYWORDS	EST.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
TITLE	1 (bases 1 to 758)			
JOURNAL	NH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .			
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Dr. Jim Lin, University of Iowa cDNA library preparation: Dr. M. Bento Soares, University of Iowa DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> This clone was contributed by the Brain Molecular Anatomy Project (BMAP) Seq primer: pYX-5.			



```

Db      123  TGTGGGACTCGCTGGCCATGATGTGCTCCTAGTCCACCAACAGCACACTCCTTCG 182
Qy      493  CTACTCCGTGTCCTGCTGGGCTTACGGCTTCTACGGGACATCAAGGACAGTGAGAA 552
Db      183  CTACTCCGTGTCCTGCTGGGCTACGGCTTCTACGGGACATCATCAGGACAGTGAGNA 242
Qy      553  GAAACGGTGGTGGGCTTCCAGATACGACTTTTCAGGTTTAAAGACCTTCTCTCCCA 612
Db      243  GAAACGGTGGTGGGCTTCCAGATACGACTTTTCAGGTTTAAAGACCTTCTCTCCCA 302
Qy      613  CCACGTCTATGAAGGACAGTGTCTTCTCTCCCTGCACACACACAGTGGGATCTCCAG 672
Db      303  CCACGTCTATGAAGGACAGTGTCTTCTCTCCCTGCACACACACAGTGGGATCTCCAG 362
Qy      673  GGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTGCAGGCAAAAGCAAGCAGCTGGA 732
Db      363  GGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTGCAGGCAAAAGCAAGCAGCTGGA 422
Qy      733  GGAGGACGAGAAAGCACTGTATGTTTGAAGCTGCGGAGGACGTGGAGGATGGCA 792
Db      423  GGAGGACGAGAAAGCACTGTATGTTTGAAGCTGCGGAGGACGTGGAGGATGGCA 482
Qy      793  AGTCGTCTGTGGGAAGTTTCTGCATCAATGCCAACAATGCCAACAATGCTCTGTGCTGCGC 851
Db      483  AGTCGTCTGTGGGAAGTTTCTGCATCAATGCCAACAATGCCAACAATGCTCTGTGCTGCGC 541

RESULT 13
BP309990 Sugano cDNA library, brain Homo sapiens EST 17-SEP-2004
LOCUS NR01216, mRNA sequence.
DEFINITION BP309990
ACCESSION BP309990
VERSION BP309990.1 GI:52238965
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.
TITLE Suzuki,Y., Yanashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES
source
1. 584
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NR01216"
/tissue_type="brain"
/clone_lib="Sugano cDNA library, brain"

ORIGIN
Query Match 53.9%; Score 527.4; DB 5; Length 584;
Best Local Similarity 94.2%; Pred. No. 3.3e-140;
Matches 582; Conservative 0; Mismatches 2; Indels 34; Gaps 2;
Qy 29 CCGTTGGAGGAAAGGACAGCGGATATGAAAGAAAGTGGCACCCTGTTTC 88
Db 1 CCGTTGGAGGAAAGGAAAGGACAGCGGATATGAAAGAAAGTGGCACCCTGTTTC 59
Qy 89 ACCTTAGCCTCCATCACCACCTGACATCATCGGTAAACAAATCTATGTTAACTATGAGAA 148
Db 60 ACCTTAGCCTCCATCACCACCTGACATCATCG----- 90
```

```

Qy      149  GTAATTACTGAACATGCTTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAA 208
Db      91  ---TTACTGAACATGCTTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAA 146
Qy      209  TACGACGCAATCCTGCTGTCGGGGAGATGTTTACGAGAGGTCGTCACCGTCTG 268
Db      147  TACGACGCAATCCTGCTGTCGGGGAGATGTTTACGAGAGGTCGTCACCGTCTG 206
Qy      269  ATTGGAGGACGACAGAGGAGCGCGGGGTGCACCAAGAACCCCGGGCTGTGCTGGTC 328
Db      207  ATTGGAGGACGACANAGGAGCGCGGGGTGCACCAAGAACCCCGGGCTGTGCTGGTC 266
Qy      329  CCCAGTACCTCCGGAATTGGAATCATTCGCCGAGGTCACCGACTCGTGTGTTACTCC 388
Db      267  CCCAGTACCTCCGGAATTGGAATCATTCGCCGAGGTCACCGACTCGTGTGTTACTCC 326
Qy      389  ACCGTGGGACACAGCGGACGAGAAACCTCGCGCTGCATATCGTTGTTGGGACTCGCTG 448
Db      327  ACCGTGGGACACAGCGGACGAGAAACCTCGCGCTGCATATCGTTGTTGGGACTCGCTG 386
Qy      449  GCCATGATGTGTCCTCAGTCCACCAACAGCACACTCTTTCGCTACTCCGTGTCCCTG 508
Db      387  GCCATGATGTGTCCTCAGTCCACCAACAGCACACTCTTTCGCTACTCCGTGTCCCTG 446
Qy      509  CTGGCTACGCTTCTACGGGACATCATCAAGGACGTAGAGAAACCGTGTGTTGGGT 568
Db      447  CTGGCTACGCTTCTACGGGACATCATCAAGGACGTAGAGAAACCGTGTGTTGGGT 506
Qy      569  CTTGCCAGATACGACTTTTCAAGGTTTAAAGACTTCTCTCCACCACTGCTATCAAGGG 628
Db      507  CTTGCCAGATACGACTTTTCAAGGTTTAAAGACTTCTCTCTCCACCACTGCTATCAAGGG 566
Qy      629  ACAGTGTCTTCTCTCCCT 646
Db      567  ACAGTGTCTTCTCTCCCT 584

RESULT 14
CFI38275
LOCUS UI-HF-BNO-anz-g-11-0-UI.r1 NIH_MGC_50 Homo sapiens cDNA clone
DEFINITION IMAGE:3094461 5', mRNA sequence.
ACCESSION CFI38275
VERSION CFI38275.1 GI:33253719
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalizaton and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRP, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Louis Staudt
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/humanfl.html
Seq primer: PYX-5.
Location/Qualifiers
1. 573
/organism="Homo sapiens"
```



Qy 876 CTGCCACTTGGGAGACGGGTCTTCTGACCTCATCCTCATCCGGAATGCTCCAGGTTCA 935  
Db |||||  
Qy 422 CTGCCCACTTGGGAGACGGGTCTTCTGACCTCATCCTCATCCGGAATGCTCCAGGTTCA 481  
Db |||||  
Qy 936 ATTTTCTGAGATTCTCATCAGGCACACCAACCGAGGACCAG 979  
Db |||||  
Qy 482 ATTTTCTGAGATTCTCATCANGCACACCAACCGAGGACCAG 525  
Db |||||

Search completed: September 6, 2005, 02:19:41  
Job time : 3590.6 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 6, 2005, 02:35:14 ; Search time 7073.33 Seconds  
(without alignments)  
3678.672 Million cell updates/sec

Title: US-10-631-958-10

Perfect score: 2888

Sequence: 1 MGATGAAPLQSLVWVKQR.....QLVRLFARGIENPKPDHS 537

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US10631956/runat.02092005.165815.4533/app\_query.fasta\_1.1941  
-DB=GenEmbl -QFMT=fastap -SURFIX=rg -MINMATCH=0.1 -LOOFC=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.coi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10631958 @CGN 1 1 11756 @runat.02092005.165815.4533 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2888	100.0	1654	9	CR456404 Homo sapi
2	2888	100.0	2042	9	HS457828 Homo sapi
3	2888	100.0	4413	6	AX457006 Sequence
4	2888	100.0	4445	9	AB079066 Homo sapi

5	2881	99.8	4432	6	ARS41900	Sequence
6	2880	99.7	4463	6	BD102675	Ceramide
7	2688.5	93.1	1772	9	BC067255	Homo sapi
8	2649.5	91.7	1840	6	AX224383	Sequence
9	2588	89.6	4171	6	BD183468	Novel gen
10	2588	89.6	4171	9	AB051433	Homo sapi
11	2465	85.4	1459	6	CQ730476	Sequence
12	2436.5	84.4	2830	10	AB079067	Mus muscu
13	1912.5	66.2	3661	10	AK129416	Gallus ga
14	1798	62.3	1450	5	CR386590	Sequence
15	1713	59.3	2494	5	BC074350	Xenopus l
16	1640.5	56.8	979	6	AX456998	Sequence
17	1495.5	51.8	1520	5	BC074110	Xenopus l
18	886.5	30.7	2333	3	AX112750	Ciona int
19	871	30.2	550	6	AX457005	Sequence
20	811	28.1	474	6	AX457001	Sequence
21	801.5	27.8	2064	6	AR509346	Sequence
22	801.5	27.8	3178	3	BT014925	Drosophil
23	763	26.4	1810	3	AY061001	Drosophil
24	710	24.6	10337	6	CQ608039	Sequence
25	710	24.6	22353	2	AC017269	Sequence
26	710	24.6	181009	3	AC007575	Drosophil
27	710	24.6	295289	3	AE003603	Drosophil
28	675	23.4	2173	6	CQ608040	Sequence
29	656	22.7	522	6	AX224385	Sequence
30	648	22.4	2257	8	AY362552	Arabidops
31	622.5	21.6	2263	8	AY035126	Arabidops
32	585.5	20.3	1084	6	AR510195	Sequence
33	578	20.0	1599	9	AY357073	Homo sapi
34	543	18.8	3039	9	BC004278	Homo sapi
35	543	18.8	3074	9	HS459H18	Homo sapi
36	541.5	18.8	70890	9	HS459H18	Human DNA
37	525	18.2	329	6	AX457002	Sequence
38	511.5	17.7	181772	2	AC007689	Homo sapi
39	498	17.2	260331	2	AC135400	Rattus no
40	450	15.6	240483	2	AC095955	Rattus no
41	437.5	15.1	564	6	AX718996	Sequence
42	428	14.8	1458	8	BT006621	Arabidops
43	428	14.8	1757	8	AY128394	Arabidops
44	421.5	14.6	2451	9	AK129976	Homo sapi
45	415	14.4	1323	5	CR760136	Xenopus t

#### ALIGNMENTS

CR456404 Homo sapiens BK29F11.1 full length open reading frame (ORF) cDNA  
clone (CDNA clone C22ORF:PGEM.BK29F11.1).

CR456404.1 GI:47678338

CR456404.1 GI:47678338

CDNA; chromosome 22; ORF.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1654)

Collins, J.E., Wright, C.L., Edwards, C.A., Davis, M.P., Grinham, J.A.,

Huckle, E.J., Beare, D.M., Mallya, M., Mokrab, Y.,

Cole, C.G., Goward, M.E., Aguado, B., Mallya, M., Mokrab, Y.,

Submitted (24-MAY-2004) Sanger Institute, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: c22@sanger.ac.uk Manuscript

Sanger Institute name: PGEM.BK29F11.1

Homo sapiens cDNA sequence. This sequence was generated as part of

The Wellcome Trust Sanger Institute program to isolate cDNA clones

representing the full length open reading frame of well annotated

protein coding genes on human chromosome 22. For more information

see <http://www.sanger.ac.uk/HGP/Chr22/>.

Location/Qualifiers

1. .1654

/organism="Homo sapiens"

#### FEATURES

source





ACCESSION	AJ457828	
VERSION	AJ457828.1	GI:20269072
KEYWORDS	lipid kinase; LK4 gene.	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1	
AUTHORS	Van Veldhoven, P.P.	
TITLE	A search for lipid kinases	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 2042)	
AUTHORS	Van Veldhoven, P.P.	
TITLE	Direct Submission	
JOURNAL	Submitted (18-APR-2002) Van Veldhoven P.P., Farmakologie, K.U.Leuven, Herestraat, B-3000 Leuven, BELGIUM	
FEATURES	Location/Qualifiers	
source	1..2042	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/chromosome="22"	
	/map="22q13"	
gene	1..2042	
	/gene="LK4"	
CDS	71..1684	
	/gene="LK4"	
	/codon_start=1	
	/product="putative lipid kinase"	
	/protein_id="CAD29884.1"	
	/db_xref="GI:20269073"	
	/db_xref="GOA:O8TCT0"	
	/db_xref="UniProt/Swiss-Prot:O8TCT0"	
	/translation="MGATGAAPLQSVLWVQRCVASEPARALLRWRSPGFCAGA PGADCSVPVSEIIATVEETDVHGKQSGKQMKPYFTVCHVKRARRHRWKWAOV TFWCEPOLCHLWLOTLEMLEKLTSPKHLVFINPFGKGQGRKIVYERKVALPFTL ASIITDIIVTHANOAKETLYEINIDKYGVCGDGMFSEVLHGLIGRTQRSAGVD QNHPRVLVPSLSRGIIPAGSDCVCSVTGTSDAETSALHIVGDSLAMDVSSVHH NSTLRYSLVLGYGYFDI IKDSEKKWGLARYDSGLKTFILSHCYEGTVSFLPA QHTVSPDRKPCRAQCFVCRSQKQLEEQKALYGLEAADEEWMQVVGQKFLAIN ATNMSCACRSPRGLSPAHLGDGSSDLIRKCSRNFRLRLRHTNQDQDFTFV EYRVKFOFTSKHMEDESDLKKEGKKRFGHICSSHPSCCTVSNSSNWCNDCGEVLHS PAIEVRVHCQLVRLPARGIEENPFDSSH"	
ORIGIN		
Alignment Scores:		
Pred. No.:	3,31e-228	Length: 2042
Score:	2888.00	Matches: 537
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	9	Gaps: 0
US-10-631-958-10 (1-537) x HSA457828 (1-2042)		
QY	1	MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrrpVallysGlnGlnArg 20
DB	71	ATGGGGGCGACGGGGCGGCGAGCGCTGCAATCGTCTGGTGGTGAAGCAGCAGCGC 130
QY	21	CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrrpArgSerProGlyPro 40
DB	131	TGCGCGGTGAGCTGAGCGCCGCGGGCTTCGCTGCGCTGGTGGCGAGCCCGGGGCC 190
QY	41	GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleIleAlaVal 60
DB	191	GGAGCGGGCGCCCGCGGCGGATGCTCTCTGTGCGCTGATCTGAGATCATCGCGCTT 250
QY	61	GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrrpGlnLysMetGluLys 80
DB	251	GAGGAAACAGACGTTTCAGGGAAACATCAAGGCAGTGGAAAAATGGCAAAATGAAAAA 310
QY	81	ProTyrAlaPheThrValHisCysVallysArgAlaArgArgHisArgTrrpAla 100

DB	311	CCTTACGCTTTTACAGTTCACTGTGTAAAGAGAGACGACGCGCACCGCTGGNAGTGGCGG 370
QY	101	GlnValThrPheTrrpCysProGluGlnLeuCysHisLeuTrrpLeuGlnThrLeuArg 120
DB	371	CAGGTGACTTTCTGGTGTCCAGAGAGCAGCTGTGTCACTTGTGGTCTGACAGACCCCTGGCG 430
QY	121	GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
DB	431	GAGATGCTGGAGAAGCTGACGTCCAGACCAAGACATTTTACTGGTATTATTCAACCCGCTTT 490
QY	141	GlyGlyLysGlyGlnGlyLysArgIleTrrpGluArgLysValAlaProLeuPheThrLeu 160
DB	491	GGAGAAAGGACCAAGCGGATATGAAAGAAAAGTGGCACCACTGTTTCACTTCA 550
QY	161	AlaSerIleThrThrAspIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180
DB	551	GCCTCCATCAGCACTGACATCATGTTACTGAAATGCTTAATCAGGCCAAGGAGACTCTG 610
QY	181	TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200
DB	611	TATGAGATTAACTAGACAAATACGACGGCATCGTCTGTGTGGCGGAGATGGTATGTTTC 670
QY	201	SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220
DB	671	AGCAGAGTGTGACGGTCTGATTTGGGAGGACGACAGGAGCGCGGGGTCCACCAAGAAC 730
QY	221	HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240
DB	731	CACGCCCGGGCTGTGTGTCCTCCAGTAGCTCCGGATTGGAATCATTTCCCGCAGGGTCA 790
QY	241	ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
DB	791	ACGAGACTCGGTGTGTACTCCACCGTGGGACCAAGCAGCAGCAAACTCCGGCGGTGCAT 850
QY	261	IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 280
DB	851	ATCGTGTGTGGGACTCGCTGGCCATGGATGTGTCTCCTCAGTCCACCAACACAGCACATC 910
QY	281	LeuArgTyrSerValSerLeuLeuGlyTrrpGlyPheTrrpGlyAspIleIleLysAspSer 300
DB	911	CTTCGTACTCGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 970
QY	301	GluLysLysArgTrrpLeuGlyLeuAlaArgTrrpAspPheSerGlyLeuLysThrPheLeu 320
DB	971	GAGAAGAAACCGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGT 1030
QY	321	SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340
DB	1031	TCCACCACTGCTGATGAAAGGACAGTGTCTTCTCTCTGCAACACACACACCGTGGGATCT 1090
QY	341	ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360
DB	1091	CAAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTGTCAGGCCAAAGCAAGCAGCAG 1150
QY	361	LeuGluGluGluGlnLysLysAlaLeuTrrpGlyLeuGluAlaAlaGluAspValGluGlu 380
DB	1151	CTGAGAGGAGCAGAGAAAGACCTGATGCTTGGAGCTTGGAGGAGCGTGGAGAGAG 1210
QY	381	TrrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400
DB	1211	TGGCAGTCTGCTGGGAGTTCCTGGCCATCAATGCCACAAACATGCTCTGTGCTTGT 1270
QY	401	ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420
DB	1271	GGCGGAGCCCAAGGGGCTCTCCCGGCTGCCACCTGGGAGAGCGGTCTTCTGACCTC 1330
QY	421	IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440
DB	1331	ATCCTCATCCGGAATGCTCCAGGTTCAAATTTTCTGAGATTTCTCATCAGCAGCACCAAC 1390
QY	441	GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 460
DB	1391	CAGCAGGACCAAGTTTGACTTCACCTTTTGTGAAAGTTTATCGCGTCAAGAAATTCAGGTTT 1450





Db	964																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
----	-----	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Alignment Scores:	3.39e-227	Length:	4432
Pred. No.:	2881.00	Matches:	536
Score:	99.81%	Conservative:	0
Percent Similarity:	99.81%	Mismatches:	1
Best Local Similarity:	99.81%	Indels:	0
Query Match:	99.76%	Gaps:	0
DB:	6		
US-10-631-958-10 (1-537) x AR541900 (1-4432)			
Qy	1	MetClyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnGlnArg	20
Db	1204	ATGGGGCGCAGCGGGCGCGGAGCCCGTCAATCCGTGCTGTGGTGAAGACAGCAGCGC	1263
Qy	21	CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTyrTrpArgSerProGlyPro	40
Db	1264	TGCGCCGTGAGCTTGGAGCCCGCGGGGCTCTGCTGCTGTGGCGGAGCCCGGGGCC	1323
Qy	41	GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleIleAlaVal	60
Db	1324	GGAGCCGGCGCCCGCGCGGATGCCCTGCTGTGCTGTATCTGAGATCATCGCGTT	1383
Qy	61	GluGluThrAspValHisGlyLysHisGlnClySerGlyLysTrpGlnLysMetGluLys	80
Db	1384	GAGGAAACAGACGTTCACGGAAACATCAAGGCAGTGGAAATGGCAAAATGGAAG	1443
Qy	81	ProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTyrLysTrpAla	100
Db	1444	CTTTACGCTTTACAGTTCACTGTGTAAGAGACACGACCGCACCGCTGGAAGTGGCG	1503
Qy	101	GlnValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg	120
Db	1504	CAGGTGACTTCTGGTGTCCAGAGGAGCAGCTGTGTCACTTGTGGCTGCAGACCTGCGG	1563
Qy	121	GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProphe	140
Db	1564	GAGATGTGGAGAAGCTGACGTCCAGACCAAGCATTTACTGGTATTATCAACCCGTTT	1623
Qy	141	GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu	160
Db	1624	GGAGAAAGGACCAAGCAGCGGATATGAAAGAAAGTGGCACCATGTTTCACTTCA	1683
Qy	161	AlaSerIleThrThrAspIleIleValThrGluHisAlaGlnLysGluThrLeu	180
Db	1684	GCCTCCATCACTGACATCATCGTTACTGAAATGCTATATCAGGCCAAGAGACTGT	1743
Qy	181	TyrGluLeuIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe	200
Db	1744	TATGAGATTAAATAGACAAATACAGCGCATCGTCTGTGTGGCGGAGATGGTATGTT	1803
Qy	201	SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn	220
Db	1804	AGCGAGGTGTGCACGGTCTGATTGGGAGGACGACAGAGGAGCGCGGGTGCACAGA	1863
Qy	221	HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer	240
Db	1864	CACCCCGGGCTGTGCTGCTGCCAGTAGCTCCCGATTCGGATTTGAATCATTTCCCG	1923
Qy	241	ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis	260
Db	1924	ACGAGCTGCGTGTGTTACTCCACCGTGGGACACGACGACGAGAAACCTCGCGCTG	1983
Qy	261	IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu	280
Db	1984	ATCGTTGTTGGGACTCGCTGCCATGGATGTGCTCTCAGTCCACCAACACAGCACATC	2043
Qy	281	LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer	300
Db	2044	CTTCGCTACTCCGTGTCCCTGCTGGGTACGGCTTCTACCGGGGACATCATCAAG	2103
Qy	301	GluLysLysArgTyrLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu	320
Db	2104	GAGAGAAACCGTGTGGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGACCTTCTC	2163

```

QY 321 SerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340
Db 2164 TCCACCACTGCTATGGAGGACAGTCTCTCTCCCTGCAACAACACACGGTGGATCT 2223
QY 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerIysGlnGln 360
Db 2224 CCAAGGGATAGGAAGCCCTGCCGGGAGGATGCTTTGTTGGAGCAAGCAAGCAGCAG 2283
QY 361 LeuGluGluGlnGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 380
Db 2284 CTGGAGGAGGAGCAGAGAAACACTGTATGTTGGAACTGCGGAGGAGCGTGGAGAG 2343
QY 381 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400
Db 2344 TGGCAAGTCGTCTGTGGGAAGTTCTGGCCATCAATGCCACAAACATGCTCTGTCTGT 2403
QY 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420
Db 2404 CGCCGGAGCCCGAGGGGGCTCTCCCGCGCTGCCCACTTGGGAGACGGGTCTTCTGACCTC 2463
QY 421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440
Db 2464 ATCCTCATCCGAATGCTCCAGGTTCATTTCTGAGATTCTCATCAGGCACACCAAC 2523
QY 441 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 460
Db 2524 CAGCAGGACCACTTTCATCTTCTGTTGAAGTTTATCGCGTCAAGAAATTCAGTTT 2583
QY 461 ThrSerLysHisMetGluAspGlnAspSerAspLeuLysGluGlyLysLysArgPhe 480
Db 2584 AGCTCGAAGCACATCGAGGATGAGACAGCGACCTCAAGAGGGGGGGAAGAGCGCTTT 2643
QY 481 GlyHisLysCysSerHisProSerCysCysCysThrValSerAsnSerSerTrpAsn 500
Db 2644 GGGCACATTTGACGAGGACCCCTCTGCTGCTGACCGCTCCCAACAGCTCTCGAAGC 2703
QY 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520
Db 2704 TCGCAGCGGGAGTCTCTGCACAGCCCTGCCATCGAGGTTCAGAGTCCACTGCCAGCTGTT 2763
QY 521 ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 537
Db 2764 CGACTCTTTGACGAGGAATTTGAAGAAATCGAAGCCAGACTCACAGC 2814

RESULT 6
LOCUS BD102675 4463 bp DNA linear PAT 27-AUG-2002
DEFINITION Ceramide kinase and DNA thereof.
ACCESSION BD102675
VERSION BD102675.1 GI:22648249
KEYWORDS WO 0196575-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 4463)
SUGIURA, M., KONO, K. and KOHAMA, T.
Ceramide kinase and DNA thereof
Patent: WO 0196575-A 1 20-DEC-2001;
SANKYO CO LTD, MASAKO SUGIURA, KEITA KONO, TAKAFUMI KOHAMA
OS Homo sapiens (human)
PN WO 0196575-A/1
PD 20-DEC-2001
PF 11-JUN-2001 WO 2001JP004889
PR 14-JUN-2000 JP 00P 178039
PI MASAKO SUGIURA, KEITA KONO, TAKAFUMI KOHAMA
PC C12N15/54, C12N9/12, C12N1/21, C12Q1/48, C07K16/40, A61K31/7125, PC
A61K48/00,
PC A61P3/04, A61P3/10, A61P7/00, A61P25/00, A61P29/00, A61P31/18, PC
A61P35/00
CC Ceramide kinase and DNA thereof
FH Key Location/Qualifiers

```

```

FEATURES             FT  CDS           (124)..(1734) .
source               1..4463
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:      4.14e-227      Length:      4463
Score:                2880.00         Matches:     535
Percent Similarity:   99.81%          Conservative: 1
Best Local Similarity: 99.63%          Mismatches:  1
Query Match:         99.72%          Indels:      0
DB:                  6                Gaps:        0

US-10-631-958-10 (1-537) x BD102675 (1-4463)
QY 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnGlnArg 20
Db 124 ATGGGGGCGAGCGGGGGCGGAGCGCTGCAATCCGTGCTGTGGTGAAGCAGCAGCGC 183
QY 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTTPTrpArgSerProGlyPro 40
Db 184 TGGCCGCTGAGCCCTGGAGCCCGCGGGCTCTGCTGCGCTGGTGGCGAGCCCGGGGCC 243
QY 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleIleAlaVal 60
Db 244 GGAGCCGGGCGCCCGCGCGGATGCTGCTCTGCTGCTGTATCTGAGATCATCGCCGTT 303
QY 61 GluGluThrAspValHisGlyLysGlnGlnGlySerGlyLysTrpGlnLysMetGluLys 80
Db 304 GAGGAAACAGACGTTACCGGGAACATCAAGGCAGTGGAAATGCGAGAAATGGAAGAAG 363
QY 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAla 100
Db 364 CCTTACGCTTTTACAGTTTCACTGTGTAAGAGAGACGACGCGCCCTGGAAGTGGCG 423
QY 101 GlnValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120
Db 424 CAGGTGACTTTCTGTTGTCAGAGAGAGCAGCTGTCTGCTGCTGGCTGCAGACCTTGG 483
QY 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
Db 484 GAGATGCTGGAGAAGCTGACGCTCCAGACCAAGCATTTACTGCTATTATCAACCCGTTT 543
QY 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
Db 544 GAGGAAAGAGACAAAGCAAGCGGATATATGAAAGAAAGTGGCACCCTGTTACCTTA 603
QY 161 AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180
Db 604 GCCTCCATCACCCTGACATCATCTGTTACTGAAATGCTATATCAGGCCAAGGAGACTCTG 663
QY 181 TyrGluLeuAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200
Db 664 TATCAGATTAACTAGACAAATACAGCGCATCTGCTGTGTCGGCGGAGATGCTATGTTTC 723
QY 201 SerGluValLeuHisGlyLysGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220
Db 724 AGCAGGTTGTCACGGTCTGATTTGGGAGGACGAGAGGAGCGCGGGGTTCAGCAGAAC 783
QY 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240
Db 784 CACCCCGGGCTGTGCTGGTCCCAGTAGCTCCGGATTGGAAATCATTTCCCGCAGGTCCA 843
QY 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
Db 844 ACGGACTCGGTGTGTTACTCCACCGTGGCACCAGCAGCAGCAAAACCTCGCGCTGCAT 903
QY 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 280
Db 904 ATCGTTGTTGGGACTCGCTGGCCATGGATGTGTCTCAGTCCAGTCCACCAACAGCACACTC 963

```



```
Qy 281 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleLeuLysAspSer 300
|||
Db 964 CTTCCGCTACTCCGTGTCCTCGCTGGCTACGGCTTCTACGGGGACATCATCAAGACAGT 1023
|||
Qy 301 GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 320
|||
Db 1024 GAGAAGAAACCGTGTGGGTCTTGCCAGATACGACTTTTTCAGGTTTAAAGACCTCTCTC 1083
|||
Qy 321 SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340
|||
Db 1084 TCCCAACCACTCTATGAAGGACAGTGTCTCTCTCCCTGCACAAACACACGGTGGGATCT 1143
|||
Qy 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360
|||
Db 1144 CCAAGGATAGAGAGCCCTGCGGGCAGGATGCTTTGTTGTCAGGCCAAAGCAACGACG 1203
|||
Qy 361 LeuGluGluGlnLysLysLeuAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 380
|||
Db 1204 CTGGAGGAGGAGCAGAAAGACACTGTATGGTTTGGAAAGCTGCGGAGGACGTGGAGGAG 1263
|||
Qy 381 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400
|||
Db 1264 TGGCAAGTCGCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCTCTGCTGTGT 1323
|||
Qy 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420
|||
Db 1324 CGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACACTTGGAGAGCGGGTCTTCTGACCTC 1383
|||
Qy 421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440
|||
Db 1384 ATCTCATCCGGAATGCTCCAAGTTCAATTTTCTGAGATTCTCATCAGGCACACCAAC 1443
|||
Qy 441 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 460
|||
Db 1444 CAGCAGGACCACTTGACTTCACCTTTGTTGAAGTTTATCGCGTCAAGAAATTCAGTTT 1503
|||
Qy 461 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyGlyLysLysArgPhe 480
|||
Db 1504 ACGTCGAGACATCGGAGATGAGGACAGCAGCCTCAAGAGGGGGGGAAGAGCGCTTT 1563
|||
Qy 481 GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn 500
|||
Db 1564 GGGCACATTTGAGCAGCAGCACCCCTCTCTGCTGTCACCGTCTCAACAGCTCTGGAAAC 1623
|||
Qy 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520
|||
Db 1624 TCGCATGGGAGGTCTCTGCACAGCCCTGCCATCGAGGTGAGAGTCCACTGCCAGCTGGTT 1683
|||
Qy 521 ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 537
|||
Db 1684 CGACTCTTGCAGGAGGAATTGAGAGAAATTCGAGACCCAGACTCACAGC 1734
|||
RESULT 7
BC067255 LOCUS 1772 bp mRNA linear PRI 06-APR-2004
DEFINITION Homo sapiens ceramide kinase, mRNA (cDNA clone IMAGE:6185601),
complete cds.
ACCESSION BC067255
VERSION BC067255.1 GI:45595582
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1772)
AUTHORS Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Alschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J.M., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
```

```
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahney,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalilus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2 (bases 1 to 1772)
Strausberg,R.
Direct Submission
Submitted (12-MAR-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapsb-x@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabh,
Parvaneh Saeedi, JR Santos, Angélique Schnerch, Ursula Skalska,
Duane Smalilus, Jeff Stott, Miranda Tsai, George Yang, Jacquie
Schein, Asim Siddiqui, Rob Holt, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 141 Row: d Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 21703365
This clone has the following problem: The cds is short compared to
the longest cds in the locus.
Location/Qualifiers
1..1772
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6185601"
/tissue_type="Peripheral Nervous System, dorsal root
ganglion"
/clone_lib="Lupski_dorsal_root_ganglion"
/lab_host="DH10B"
/notes="vector: pCMV-SPORT6"
1..1772
/gene="CERK"
/notes="synonyms: KIAA1646, hCERK, dA59H18.3,
DKFZp344E0211, FLJ21430, FLJ23239, dA59H18.2, LK4"
/db_xref="LocusID:64781"
579..1598
/gene="CERK"
/product_start=1
/product="CERK protein"
/protein_id="AAH67255.1"
/db_xref="GI:45595583"
/db_xref="LocusID:64781"
/translaton="MFSEVLHGLIGRTORSAGVNDHPRAVLVPSSLRIGTIPAGSTD
CVCYSTVGTSDAETSAUHI VVGSLANDMVSSVHNSILLRYSVLLGYGFGDILKDS
EKRWLGLARYDFSLGKLTGTVSHHCYGTGTVFLPAQHTVSGPRDRKPRAGCFVCRQSK
QOLEEEOKALYGLAEAEVDEVMQVCGKFLAINATNMSCACRRSPRGLSPAHLGDG
```



SSDLILIRKSRFNLRLIHTNOODQDFTFVEVVRVKKFQFTSKHWEDESDLKE  
GCKKRFHCHSHSPSCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIENPK  
PUSHS"

## ORIGIN

Alignment Scores:

Pred. No.: 7,68e-212 Length: 1772  
Score: 2688.50 Matches: 511  
Percent Similarity: 94.98% Conservative: 0  
Best Local Similarity: 94.98% Mismatches: 1  
Query Match: 93.09% Indels: 27  
DB: 9 Gaps: 2

US-10-631-958-10 (1-537) x BC067255 (1-1772)

Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnGlnArg 20  
Db 60 ATGGGGCGACGGGGCGCGGAGCGCTGCAATCGTGCTGTGGTGAAGCAGCAGCGC 119  
Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40  
Db 120 TCGCGCGTGAGCTGGAGCGCGCGGGCTCTGCTGCGCTGCTGGCGGAGCGCGGGGCC 179  
Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal 60  
Db 180 GGAGCGCGCGCGCGCGCGGATCGCTCTGTGCTGTATCTGAGATCATCGCGCTT 239  
Qy 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80  
Db 240 GAGGAAACAGACGTTTACGGGAAACATCAAGCAGTGGGAAATGCGAGAAATGGAAAG 299  
Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrpLysTrpAla 100  
Db 300 CCTTACGCTTTTACAGTTTCACTGTGTAAGAGAGCAGCAGCAGCGCTGGAAGTGGCG 359  
Qy 101 GlnValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120  
Db 360 CAGTGTACTTCTGTGTGTCAGAGAGCAGCTGTGTCACTTGTGCTGCAGACCCCTGGCG 419  
Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140  
Db 420 GAGATGCTGGAGAGCTCAGCTCCAGACCAAGCAATTTACTGGTATTATCAAC----- 473  
Qy 141 GlyGlyLys-GlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLe 160  
Db 474 -----AAAAGGCAAGGCAAGCGGATATGAAAGAAAGTGGCACCACCTGTTCACTT 527  
Qy 160 uAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLe 180  
Db 528 AGCTTCCATCACCCTGACATCATCGC----- 554  
Qy 180 uTyrGluIleAsnIleAspLysTyrAspGlyLeValCysValGlyGlyAspGlyMetPh 200  
Db 555 -----ATGCTGTGTGTCGGCGAGATGTAATGTT 583  
Qy 200 eSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAs 220  
Db 584 CAGCGAGTGTGTCAGCTCTGATTGGGAGGACGAGAGGCGCGCGGGTCGACACGAA 643  
Qy 220 nHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAlaGlySe 240  
Db 644 CCACCCCGGCTGTGCTGGTCCCGAGTAGCTCCGGATTGGAATCATTCCTCCGAGGGTC 703  
Qy 240 rThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHi 260  
Db 704 AACGGAGCTGTGTGTTACTCCAGCTGGGCAACGAGCGAGCAAACTCGGCGCTGCA 763  
Qy 260 sIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLe 280  
Db 764 TATCTGTTGGGACTCGCTGGCCATGGATGTGTCTCAGTCCACCAACACAGCACACT 823  
Qy 280 uLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSe 300

Db 824 CCTCGCTACTCCGTGTCCCTCGTGGGCTACGGCTTCTACGGGGGACATCATCAAGACAG 883  
Qy 300 rGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLe 320  
Db 884 TGAGAAGAAACGGTGGTGGTCTTGCAGATACGACTTTTCAGGTTTAAAGACCTTCTCT 943  
Qy 320 uSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySe 340  
Db 944 CTCCACCACTGCTATGAAGGAGCAGTGTCTCTCCCTGTCACAAACACACGCTGGATC 1003  
Qy 340 rProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnG 360  
Db 1004 TCCAAGGATAGGAAGCCCTCCGGCGAGGATGCTTTGTTTCAGCGCAAGCAACAGCA 1063  
Qy 360 nLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluG 380  
Db 1064 GCTGGAGGAGGAGCAGAGAAAGCACGTATGTTTGGAGCTGCGGAGGAGCTGGAGGA 1123  
Qy 380 uTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCy 400  
Db 1124 GTGGCAAGTCTCTGTGGGAAAGTTTCTGGCCATCAATGCCCAACATGTCTGTGCTTG 1183  
Qy 400 sArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLe 420  
Db 1184 TCGCGGAGGCCCGAGGGCCCTCTCCCGCTGCCACTTTGGGAGACGGGTCTTCTGACCT 1243  
Qy 420 uIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAs 440  
Db 1244 CATCTCATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAA 1303  
Qy 440 nGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPh 460  
Db 1304 CCAGCAGGACCACTTTCAGTTTCTTGTGTAAGTTTATCGCGTCAAGAAATTCAGTT 1363  
Qy 460 eThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPh 480  
Db 1364 TACGTGAGACACATGAGGATGAGCAGCGACTCAAGAGGGGGGAGAGACGCTT 1423  
Qy 480 eGlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAs 500  
Db 1424 TGGGCACATTTGACAGGACCACTCTCTGCTGTCACCGCTCTCCAACAGCTCTCTGGA 1483  
Qy 500 nCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVa 520  
Db 1484 CTGCGAGCGGAGGTCCTGCACAGCCCTGCTCCATCGAGTCAGAGTCCACTGCCAGTGT 1543  
Qy 520 lArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 537  
Db 1544 TCGACTCTTTGACGAGGATTTGAAGAAATCCGAAGCCAGACTCACACAGC 1595

RESULT 8  
AX224383  
LOCUS AX224383 1840 bp DNA linear PAT 10-SEP-2001  
DEFINITION Sequence 5 from Patent WO0160990.  
ACCESSION AX224383  
VERSION AX224383.1 GI:15554633  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Rastelli, L.  
TITLE Novel sphingosine kinases  
JOURNAL Patent: WO 0160990-A 5 23-AUG-2001;  
Curagen Corporation (US); GENENTECH, INC. (US)  
FEATURES  
Location/Qualifiers  
1..1840  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

## Alignment Scores:

Pred. No.: 1.32e-208 Length: 1840  
 Score: 2649.50 Matches: 495  
 Percent Similarity: 97.83% Conservativity: 1  
 Best Local Similarity: 97.63% Mismatches: 0  
 Query Match: 91.74% Indels: 11  
 DB: 6 Gaps: 1

US-10-631-958-10 (1-537) x AX224383 (1-1840)

Qy 42 AlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleLeuAlaValGlu 61  
 Db 1 GCCGCGCCCCCGCGCGGATGCCCTGCTGCTGCTGCTATCTAGATCATCGCGTTGAG 60  
 Qy 62 GluThrAspValHisGlyLeuHisGlnGlySerGlyValTrpGlnLysMetGluLysPro 81  
 Db 61 GAAACAGACGTTACCGGGAACATCAAGCGCAGTGGAAATGGCAGAAATGGAAAGACCT 120  
 Qy 82 TyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrpLysTrpAlaGln 101  
 Db 121 TAGCTTTTACAGTTTCACTGTGTAAAGAGACGACGCGCCGCTGGAAGTGGCGCAG 180  
 Qy 102 ValThrPheTrpCysProGluGlnLeuGlnCysHisLeuTrpLeuGlnThrLeuArgGlu 121  
 Db 181 GTGACTTTCTGGTGTCCAGAGGACAGCTGTGTCACTTGTGGCTGCAGACCTCGCGGAG 240  
 Qy 122 MetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPheGly 141  
 Db 241 ATGCTGGAGAACGTGACGTCCAGACCAAGCAATTTACTGTGTATTTATCAACCCGTTTGA 300  
 Qy 142 GlyLysGlyGlnGlnLysArgIleTyrGluArgLysValAlaProLeuPheThrLeuAla 161  
 Db 301 GGAAGAGACAAAGCGGATATATGAAAGAAAAGTGGCACCACCTGTTCACTTAGCC 360  
 Qy 162 SerIleThrThrAspIleIle-----ValThr 170  
 Db 361 TCCATCACCATGACATCATCGGTAAACAAATTCATGTAACTATGTAGAAGTAATTACT 420  
 Qy 171 GluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGly 190  
 Db 421 GAACATGCTAATCAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATAGCAGGC 480  
 Qy 191 IleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArg 210  
 Db 481 ATCGTCTGTGCGCGCAGATGTTATGTTTCAGCGAGTGTCTGCACCGTCTGATTGGGAGG 540  
 Qy 211 ThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSer 230  
 Db 541 ACCGAGAGGAGCGCGGGGTGACCAAGAACCCCGCGGCTGTGCTGGTCCCGAGTAGC 600  
 Qy 231 LeuArgIleGlyIleIleProAlaGlySerThrAspCysValCysTyrSerThrValGly 250  
 Db 601 CTCGGATTGGATTATTCGCGCAGGTCAACGAGACTGCGTGTGTTACTCCACGTTGGC 660  
 Qy 251 ThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAsp 270  
 Db 661 ACCAGCGACGAGAAACCTCGCGGCTGCATATCGTTTGGGACTCGCTGGCCATGGAT 720  
 Qy 271 ValSerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyr 290  
 Db 721 GTGTCTTCAGTCCACCAACAGACACACTCTCTTCGCTACTCCGTTGCTCTGCTGGGCTAC 780  
 Qy 291 GlyPheTyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArg 310  
 Db 781 GGCTTTCTACGGGGACATCATCAAGACAGTGTAGAGAAACGGTGTGGTCTTGGCCAGA 840  
 Qy 311 TyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSer 330  
 Db 841 TACGACTTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGCTGTCC 900  
 Qy 331 PheLeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGly 350  
 Db 901 TTCTCTCCCTGCACAAACACACGGTGGGATCTCCAAAGGGATAGGAAGCCCTCGCGGACGA 960

Qy 351 CysPheValCysArgGlnSerLysGlnGlnLeuGluGlnGluGlnLysLysAlaLeuTyr 370  
 Db 961 TGCTTTGTTTGCAGGCAAGCAAGCAGCAGCTGGAGGAGGAGCAGAAAGGACTGTAT 1020  
 Qy 371 GlyLeuGlnAlaAlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAla 390  
 Db 1021 GGTTTGAAGCTCGGAGGACGTGGAGGAGTGGCAAGTCTGTGTGGAAAGTTTCTGGCC 1080  
 Qy 391 IleAsnAlaThrAsnMetSerCysAlaCysArgSerProArgGlyLeuSerProAla 410  
 Db 1081 ATCAATGCCCAACAACATGCTCTGTGTGTGCGCGGAGCCCGAGGCGCTCTCCCGGCT 1140  
 Qy 411 AlaHisLeuGlyAspGlySerSerAspLeuIleLeuIleArgLysCysSerArgPheAsn 430  
 Db 1141 GCCACTTGGGAGACGGGTCTTCTGACCTCATCTCATCCGGAATGCTCCAGGTTCAAT 1200  
 Qy 431 PheLeuArgPheLeuIleArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheVal 450  
 Db 1201 TTTCTGAGATTTCTCATCAGGCACACCAACAGCAGGACCAGTTTGACTTCACTTTTGT 1260  
 Qy 451 GluValTyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSer 470  
 Db 1261 GAAGTTTATCGCTCAAGAAATTCAGTTTACGTGCAAGCACATGAGGAGATGAGGACG 1320  
 Qy 471 AspLeuLysGluGlyGlyLysLysArgPheGlyHisIleCysSerSerHisProSerCys 490  
 Db 1321 GACTCAAGGAGGGGGGAGAGAGCGCTTGGGACATTTGACAGCAGCCACCTCTCTGC 1380  
 Qy 491 CysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAla 510  
 Db 1381 TGCTGACCGCTCTCCAACAGCTCTCGAACTGCGACGGGGAGGTCTTGCAACAGCCCTGCC 1440  
 Qy 511 IleGluValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsn 530  
 Db 1441 ATCGAGGTCAGGTTCACCTGCCAGCTGGTTCGACTCTTTTGACGAGGAATGAAGAAT 1500  
 Qy 531 ProLysProAspSerHisSer 537  
 Db 1501 CCGAGCCAGACTCACACAGC 1521

RESULT 9  
 BD183468  
 LOCUS 4171 bp DNA linear PAT 17-JUN-2003  
 DEFINITION Novel genes and proteins encoded by the genes.  
 ACCESSION BD183468  
 VERSION BD183468.1 GI:31875668  
 KEYWORDS JP 2002345492-A/181.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 4171)  
 AUTHORS Ohara, O., Nagase, T. and Nakajima, D.  
 TITLE Novel genes and proteins encoded by the genes  
 JOURNAL Patent: JP 2002345492-A 181 03-DEC-2002;  
 KAZUSA DNA RESEARCH INSTITUTE  
 COMMENT OS Homo sapiens (human)  
 PN JP 2002345492-A/181  
 PD 03-DEC-2002  
 PF 26-FEB-2002 JP 2002049009  
 PI OHSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA  
 PC C12N15/09, C07K14/47//A61K31/711, A61K38/00, A61K48/00, A61P25/00,  
 A61P25/14,  
 PC A61P25/18, A61P35/00, C12N15/00, A61K37/02  
 CC Novel genes and proteins encoded by the genes FH Key  
 Localization/Qualifiers  
 CDS (2)..(1444).  
 FEATURES  
 source 1..4171  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

## ORIGIN

## Alignment Scores:

Pred. No.: 4,466-203 Length: 4171  
Score: 2588.00 Matches: 481  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 89.61% Indels: 0  
DB: 6 Gaps: 0

US-10-631-958-10 (1-537) x BD183468 (1-4171)

QY 57 llelleAlaValGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGln 76  
DB 2 ATCATCGCGGTTGAGGAAACACAGCTTCACGGGAAACATCAAGGCAGTGGAAATGGCAG 61  
QY 77 LysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgHisArg 96  
DB 62 AANATGGAAAGCCTTACGCTTTTACAGTTCACTGTGTAAAGAGACGACGCGCACCGC 121  
QY 97 TrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuLysHisLeuTrpLeu 116  
DB 122 TGGAAAGTGGCGCAGGTACTTTCTGTGTCCAGAGGAGCAGCTGTGTCACTTTGTGGCTG 181  
QY 117 GlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuValPhe 136  
DB 182 CAGACCCCTGCGGGAGATGCTGGAGAAGCTGACGCTCCAGACCAAGCATTTTACTGGTATTT 241  
QY 137 lleAenProPheGlyLysGlyGlnGlyLysArgGlyLeuTyrGluArgLysValAlaPro 156  
DB 242 ATCAACCCGTTTGGAGGAAAGGACAGGCAAGCGGATATATGAAGAAAGTGGCACCA 301  
QY 157 LeuPheThrLeuAlaSerIleThrAspIlelleValThrGluHisAlaAasnGlnAla 176  
DB 302 CTGTTCACTTACGCTCCATCACCACCTGACATCATCGTTACTGAACATGCTAATCAGGCC 361  
QY 177 LysGluThrLeuTyrGluIleAenIleAspLysTyrAspGlyIleValCysValGlyGly 196  
DB 362 AAGGAGACTGTATGAGATTAAACATAGACAAATACGACGCGCATCGTGTGTGCGCGGA 421  
QY 197 AspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGly 216  
DB 422 GATGGTATGTTGAGGAGGTGCTGACCGTCTGATTTGGAGGACCAAGAGGCCCGGG 481  
QY 217 ValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIle 236  
DB 482 GTCCAGCAGAACCAACCCCGGGCTGTGTGTGTCCTCCAGTAGCTCCGGATTGGATCAT 541  
QY 237 ProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThr 256  
DB 542 CCGCGCAGGGTCAACGAGACTGCGTGTGTACTCCACCGTGGGCACCGCAGCGCAGAAACC 601  
QY 257 SerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHis 276  
DB 602 TCGGCGCTGCATATCGTTGTGGGACATCGTGGCCATGGATGTCTCTCAGTCCACCAC 661  
QY 277 AenSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIle 296  
DB 662 AACACACACTCTTCGCTACTCCGTGTCCTGCTGGCTACGGTCTTACGGGACATC 721  
QY 297 lleLysAspSerGluLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeu 316  
DB 722 ATCAAGGACAGTGAGAAGAAACGGTGTGGTCTTGGCAGATACGACTTTTCAGTTTA 781  
QY 317 LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis 336  
DB 782 AAGACCTTCTCTCCACCACCTGCTATGAAGGAGCAGTGTCTCTCTCCCTGCACAAACAC 841  
QY 337 ThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGln 356  
DB 842 ACGTGGGATCTCCAAGGAGATAGGAGCCCTCCCGGCGCAGGATGTTGTTTGGAGGCA 901  
QY 357 SerLysGlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGlu 376

DB 902 AGCAAGCAGCAGCTGGAGGAGGAGGAGAAAGAACACTGTATGTTTGGAAAGCTGCGGAG 961  
QY 377 AspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMet 396  
DB 962 GACGTGGAGGAGTGGCAAGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATG 1021  
QY 397 SerCysAlaCysArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGly 416  
DB 1022 TCCTGTGCTTGTGCGCGGAGGCCCGCAGGGCCCTCTCCCGGCTGCGCACTTGGGAGCGG 1081  
QY 417 SerSerAspLeuIleLeuIleArgLysCysSerArgPheAenPheLeuArgPheLeuIle 436  
DB 1082 TCTTCTGACCTCATCTCTCATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATC 1141  
QY 437 ArgHisThrAenGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLys 456  
DB 1142 AGGCACCAACACAGCAGGACCACTTTCAGTTCTCACTTTTGTGAAGTTTATCGCTCAAG 1201  
QY 457 LysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyGly 476  
DB 1202 AATTTCCAGTTTACGTTGAGACATGAGGATGAGGACAGCAGCCTCAAGGAGGGGG 1261  
QY 477 LysLysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSerAen 496  
DB 1262 AAGAAGCGCTTGTGGCACATTTGACGAGCAGCCCTCTCTGCTGTCACCGTCTCCAAC 1321  
QY 497 SerSerTrpAenCysAspGlyValLeuHisSerProAlaIleGluValArgValHis 516  
DB 1322 AGCTCTCGAACTCGGACGGGAGGTCTGACAGAGCCTGCGCATCGAGTCCAGTCCAC 1381  
QY 517 CysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAenProLysProAspSerHis 536  
DB 1382 TGCCAGCTGGTTCGACTCTTTGCGACGAGGAATGAAGAGAATCGAAGCCAGACTCACAC 1441  
QY 537 Ser 537  
DB 1442 AGC 1444  
RESULT 10  
AB051433  
LOCUS AB051433 4171 bp mRNA linear PRI 06-OCT-2001  
DEFINITION Homo sapiens mRNA for KIAA1646 protein, partial cds.  
ACCESSION AB051433  
VERSION AB051433.1 GI:13359166  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 Hirosawa,M., Nagase,T., Murahashi,Y., Kikuno,R. and Ohara,O.  
Identification of novel transcribed sequences on human chromosome  
22 by expressed sequence tag mapping  
DNA Res. 8 (1), 1-9 (2001)  
MEDLINE 21156230  
PUBMED 11258795  
REFERENCE 2 (bases 1 to 4171)  
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.  
DIRECT SUBMISSION  
TITLE Direct Submission  
JOURNAL Submitted (22-NOV-2000) Osamu Ohara, Kazusa DNA Research Institute,  
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba  
292-0812, Japan (E-mail:cdnaifookazusa.or.jp,  
URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,  
Fax:81-438-52-3914)  
FEATURES  
Location/Qualifiers  
1. .4171  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="hk01650"  
/note="vector:pBluescriptII SK plus"  
1. .4171  
gene

CDs		/gene="KIAA1646"		Qy		277		AsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIle		296	
		<1..1447		Db		662		AACAGCACACTCTTGGTACTCCGTGCTGCGGCTACGGCTTCTACGGGACATC		721	
		/gene="KIAA1646"		Qy		297		IleIysAspSerGluIysIysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeu		316	
		/note="Start codon is not identified."		Db		722		ATCAAGCAGACTGAGAAGAACCGTGGTGGGCTTGGCCAGATACGACTTTTCAGGTTA		781	
		/codon_start=2		Qy		317		LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHis		336	
		/product="KIAA1646 protein"		Db		782		AAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTTCTCCCTGCACACAC		841	
		/protein_id="BAB33316.1"		Qy		337		ThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGln		356	
		/db_xref="GI:13359167"		Db		842		ACGGTGGGATCTCCAAAGGATAGGAAGCCCTGCGGGCAGATGCTTTGTTGCGAGCAA		901	
		/translators="IIAETEDVHGKHQSGKQMKPKYAFVTHCVKRRRRHWKA		Qy		357		SerIysGlnGlnLeuGluGluGlnIysLysAlaLeuTyrGlyLeuGluAlaAlaGlu		376	
		QVTFWCPBEQLCHLWLTQREMLEKLTSRPKHLVFNPPGGKGGKRIYERKVAPLF		Db		902		AGCAAGCAGCAGCTGGAGGAGGAGCAGAAAGACACTGTATGTTTGGGAAGCTGCGGAG		961	
		TLASITDIIIVTEHANAKETLYEINDIKYDGVICVGGDGMFSEVLHGLIGRTQRSAG		Qy		377		AspValGluGluTyrGlnValValCysGlyLysPheLeuAlaIleAenAlaThrAsnMet		396	
		VDQNHPRALVLSLGIIPAGSTDCVYSTGVTSDAETSAHLHIVVGDLSLAMDSSV		Db		962		GACGTGGAGGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATG		1021	
		HNSTLLARYSVSLGYGDI IKSEKRWMLGLARYDFSLKTFILSHCYEGTVSFL		Qy		397		SerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGly		416	
		PAQHTGSPRDKPCRCGFCVROSQKQLEEKALYGLEAAEDVEEVMQVCGKFLA		Db		1022		TCCTGTGTTGTGCGGAGCCCGAGGGGCTCTCCCGGCTGCCACTTTGGGAGACGGG		1081	
		INATNMSCACRRSPRGLSPAHLGDGSSDLILIRKSRFNFLRLIRHTNQDQDFPT		Qy		417		SerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIle		436	
		FVEYRVKQFQTSKMEDESDLEKGGKRGHICSSHPSCCCTVSNSNSNCDEVL		Db		1082		TCYTTGACCTCATCTCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTTCCTATC		1141	
		HSPAIEVRVHCQLVRLFARGIEENPKPDSHS"		Qy		437		ArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLys		456	
				Db		1142		AGGCACACCAACACGACGAGGACAGTTGACTTCACTTTTGTGAGTTTATCGGTCAG		1201	
				Qy		457		LysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGly		476	
				Db		1202		AAATTCCAGTTTACGTCGAAGCACATGAGGATGAGGACAGCACCTCAAGAGGGGGGG		1261	
				Qy		477		LysIysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSerAsn		496	
				Db		1262		AAGAAGCGCTTTGGGCACATTTGCAGCAGCACCCCTCTCTGCTGCTGACCGTCCAAC		1321	
				Qy		497		SerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHis		516	
				Db		1322		AGTCTGGAACTCGGAGCGGGAGGCTCTGCACAGCCCTGCCATCGAGGTCCAC		1381	
				Qy		517		CysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHis		536	
				Db		1382		TGCCAGCTGGTTCGACTCTTTGCACGAGGAATTCGAAGAAATTCGAAGCCAGACTCACAC		1441	
				Qy		537		Ser 537			
				Db		1442		AGC 1444			
				RESULT 11							
				LOCUS		CQ730476		1459 bp		DNA	
				DEFINITION		Sequence 16410 from Patent WO02068579.				linear	
				ACCESSION		CQ730476					
				VERSION		CQ730476.1		GI:42304409			
				KEYWORDS							
				SOURCE		Homo sapiens (human)					
				ORGANISM		Homo sapiens					
				REFERENCE		1		Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.			
				AUTHORS							
				TITLE							
				JOURNAL							

## ORIGIN

Alignment Scores:			
Pred. No.:	4.46e-203	Length:	4171
Score:	2588.00	Matches:	481
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	89.61%	Indels:	0
DB:	9	Gaps:	0
US-10-631-958-10 (1-537) x AB051433 (1-4171)			
Qy	57	IleIleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTyrGln	76
Db	2	ATCATCGCGTGGAGAAACAGACGTTACGGGAACATCAAGGACGTGGAAAAATGGCAG	61
Qy	77	LysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArg	96
Db	62	AAAAATGGAAAAGCCTTACGCTTTTACAGTTTCACTGTGTAAAGAGACGACGACGCGC	121
Qy	97	TrpLysTrpAlaGlnValThrPheTyrCysProGluGluGlnLeuCysHisLeuTyrLeu	116
Db	122	TGGAAAGTGGGCGCAGGTGACTTTCTGGTGTCCAGAGGAGCAGCTGTGTCACTTGTGGCTG	181
Qy	117	GlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPhe	136
Db	182	CAGACCTTCGCGGAGATGCTCGAGAAGCTGACGCTCCAGACCAAGCATTTTACTGGTATT	241
Qy	137	IleAsnProPheGlyGlyLysGlnGlyLysArgIleTyrGluArgLysValAlaPro	156
Db	242	ATCAACCCGTTTGGAGGAAAGGACAGGCAAGCGGATATAGAAAGAAAGTGGCACCA	301
Qy	157	LeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAla	176
Db	302	CTGTTCACTTAGCTTCATCCACCTGACATCATCTGTTACTGAACATGCTAATCAGGCC	361
Qy	177	LysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGly	196
Db	362	AAGGAGACTGTGTATGAGATTAAACATAGACAAATACGACGCGCATCGTGTGTGCGGGA	421
Qy	197	AspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGly	216
Db	422	GATGTAATGTTACAGCGAGTGTGCACGCTGTGATTTGGAGGAGCAGAGAGGCGCGGG	481
Qy	217	ValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIle	236
Db	482	GTCGACCAAGAACACCCCGGGCTGTGCTGTGCTCCAGTAGCCTCCGGATTGGAAATCAT	541
Qy	237	ProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThr	256
Db	542	CCCGCAGGGTCAACGAGCTGCGTGTGTACTCCACCGTGGGCACGACGACGAGAAACC	601
Qy	257	SerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerValHisHis	276
Db	602	TCGGCGCTGCATATCGTTTGGGAGCTCGCTGGCCATGGATGTGCTCTCAGTCCACCAC	661

Qy	377	AspValGluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMet	396
Db	962	GACGTGGAGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAACATG	1021
Qy	397	SerCysAlaCysArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGly	416
Db	1022	TCTGTGCTTGTGCGCGGAGCCCGAGGGGCTCTCCCCGGCTGCCCATCTTGGAGACGGG	1081
Qy	417	SerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIle	436
Db	1082	TCTTCTGACCTCATCTCTATCCGGAAATGCTCCAGGTTCAATTTCTGAGATTTCTCATC	1141
Qy	437	ArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLys	456
Db	1142	AGGCACCAACACAGCAGGACCACTTTGACTTTCATCTTTTGTGAAGTTTATCGGCTCAAG	1201
Qy	457	LysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyGly	476
Db	1202	AAATTCAGTTTACGTCGAAGCACATGGAGGATGAGGACGCGACCTCAAGGAGGGGGG	1261
Qy	477	LysLysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSerAsn	496
Db	1262	AGAAAGCGCTTTGGGCACATTTGCAGACGCCACCCCTCTGCTGCTGCACCGTCTCCAC	1321
Qy	497	SerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHis	516
Db	1322	AGCTCTCGGAACCTGCGAGGGGAGGCTCTGCACACGCCCTGCCATCGAGGTCCAGATCCAC	1381
Qy	517	CysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHis	536
Db	1382	TGCCAGCTGGTTCGACTCTTTGCAGAGGAATTGAAGAGAATCCGAAGCCAGACTCACAC	1441
Qy	537	Ser 537	
Db	1442	AGC 1444	
RESULT 11			
LOCUS	Q0730476	1459 bp	DNA linear PAT 03-FEB-2004
DEFINITION	Sequence 16410 from Patent WO02069579.		
ACCESSION	Q0730476		
VERSION	Q0730476.1	GI:42304409	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	1		
AUTHORS	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.		
TITLE	Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof		
JOURNAL	Patent: WO 02068579-A 16410 06-SEP-2002;		

```

PE Corporation (NY) (US)
Location/Qualifiers
1. .i459
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 1.59e-193 Length: 1459
Score: 2465.00 Matches: 470
Percent Similarity: 87.71% Conservative: 1
Best Local Similarity: 87.52% Mismatches: 6
Query Match: 85.35% Indels: 61
DB: 6 Gaps: 3

US-10-631-958-10 (1-537) x CQ730476 (1-1459)

Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrrpValLysGlnArg 20
Db 27 ATGGGGGCGACGGGGCGCGGAGCGCTGCAATCCGTGCTGTGGTGAAGCAGCAGCGC 86
Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrrpArgSerProGlyPro 40
Db 87 TGGCGCAA----- 94
Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal 60
Db 95 GGACGTGGTGT---TTTACAGATCCCTGCTCTGTGCTGTATCTGAGATCATCGCGGT 151
Qy 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrrpGlnLysMetGluLys 80
Db 152 GAGGAACACAGACGTTTCAGGGGAAACATCAAGGCAGTGGAAAAATGGCAAAATGGAAAG 211
Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrrpLysTrpAla 100
Db 212 CTTTACGCTTTTACAGTTTACATGTGTAAAGAGACGACGCGCCGCTGGAAAGTGGCG 271
Qy 101 GlnValThrPheTrrpCysProGluGlnLeuGlnCysHisLeuTrrpLeuGlnThrLeuArg 120
Db 272 CAGGTGACTTCTCTGTGTCCAGAGGAGCAGCTGTGTCTCACTTGTGGCTGCAGACCCCTGGG 331
Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
Db 332 GAGATGCTGGAGAGCTG----- 349
Qy 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
Db 349 ----- 349
Qy 161 AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180
Db 350 -----ATTACTGAACATGCTAATCAGGCCAAGGAGACTCTG 385
Qy 181 TyrGluIleAsnIleAspLysTrrpAspGlyLysValGlyCysValGlyAspGlyMetPhe 200
Db 386 TATGAGATTAAACATAGACAAATPACACGCGCATCGTCTGTCTGGCGGAGATGTTATGTTTC 445
Qy 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220
Db 446 ACAGAGGTGCTGCACGGTCTGATTGGAGGACGACGAGGAGCGCGGGGTGCACAGAAC 505
Qy 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240
Db 506 CACCCCGGGCTGTCTGTTCCAGTAGCTCCGATTGGATTCATTTCCCGCAGGTCA 565
Qy 241 ThrAspCysValCysTrrpSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
Db 566 ACAGACTGCTGTGTACTTCCACCGTGGGCACGACGACGAGAAACCTCGCGCTGCAT 625
Qy 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 280
Db 626 ATCGTTTGTGGGACTCGCTGGCCATGATGTCTCAGTCCACCAACAGCAGCACTC 685

```

---

```

Qy 281 LeuArgTyrSerValSerLeuLeuGlyTrrpGlyPheTrrpGlyAspIleIleLysAspSer 300
Db 686 CTTTCGTACTCGGTGTCCTGCTGGCTACGGCTTCTACGGGACATCATCAAGGACAGT 745
Qy 301 GluLysLysArgTrrpLeuGlyLeuAlaArgTrrpAspPheSerGlyLeuLysThrPheLeu 320
Db 746 GAGAAGAAACGGTGGTGGTCTTGGCAGATACACATTTTCAGGTTTAAAGACCTTCCTC 805
Qy 321 SerHisHisCysTrrpGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340
Db 806 TCCACCACTGCTATGAAAGGACAGTGTCTTCTCTCTGCAACAACACACCGTGGATCT 865
Qy 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360
Db 866 CCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTTCAGGCAAGCAAGCAGCAG 925
Qy 361 LeuGluGluGlnLysLysAlaLeuTrrpGlyLeuGluAlaAlaGluAspValGluGlu 380
Db 926 CTGAGAGGAGGACAGAAAGACATGTATGTTTGAAGCTGCGGAGGACGTGGAGGAG 985
Qy 381 TrrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400
Db 986 TGGCAAGTGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCTCTGTCTGT 1045
Qy 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420
Db 1046 CGCGGAGCCCGAGGGGCTCTCCCGGCTGCCCATCTGGGAGACGGGTCTTCTGACCTC 1105
Qy 421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440
Db 1106 ATCTCTATCCGGAATGCTCCAGGTTCATTTCTGAGATTTCTCATCAGCACACCAAC 1165
Qy 441 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 460
Db 1166 CAGCAGACCCAGTTTGACTTCACTTTTGTGAAGTTTATCCGCTCAAGAAATTCAGTTT 1225
Qy 461 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe 480
Db 1226 ACGTGGAAGCACATGGAGGATGAGCAGCAGCCTCAAGGAGGGGGGAGAGCGCTTT 1285
Qy 481 GlyHisIleCysSerSerHisProSerCysCysCysThrValSerAsnSerSerTrpAsn 500
Db 1286 GGGCACATTTGCACAGCCACCCCTCTCTGTCTGTCACCCGTCCTCAACAGCTCTCGAAC 1345
Qy 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520
Db 1346 TGCACCGGGGAGGTCTTCACAGCCCTGCCATCCAGGTTCAGATCCACTGCCAGCTGGTT 1405
Qy 521 ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 537
Db 1406 CGACTCTTTGCACGAGGAATTGAAGAGAAATCCGAAGCCAGACTCACAGC 1456

RESULT 12
AB079067 2830 bp mRNA linear ROD 27-JUN-2002
LOCUS Mus musculus cerk mRNA for ceramide kinases, complete cds.
DEFINITION
AB079067
ACCESSION
AB079067.1 GI:21624341
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Sugiura, M., Kono, K., Liu, H., Shimizu, T., Minekura, H.,
Spiegel, S., and Kohama, T.
Ceramide Kinase, a Novel Lipid Kinase. MOLECULAR CLONING AND
FUNCTIONAL CHARACTERIZATION
J. Biol. Chem. 277 (26), 23294-23300 (2002)
JOURNAL
11956206
PUBMED
2 (bases 1 to 2830)
REFERENCE
Sugiura, M., Kono, K., Shimizu, T., Minekura, H., Spiegel, S. and
AUTHORS

```

Kohama, T.  
 Direct Submission  
 Submitted (29-JAN-2002) Masako Sugiura, Sankyo Co., Ltd.,  
 Pharmacology and Molecular Biology Research Laboratories; 2-58  
 Hiromachi 1-chome, Shinagawa-ku, Tokyo 140-8710, Japan  
 (E-mail: msugiura@shina.sankyo.co.jp, Tel: 81-3-3492-3131,  
 Fax: 81-3-5436-8565)  
 Location/Qualifiers  
 1. .2830  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 1. .2830  
 /gene="cerk"  
 1. .1596  
 /gene="cerk"  
 /codon\_start=1  
 /product="ceramide kinases"  
 /protein\_id="BAC01155.1"  
 /db\_xref="GI:21624342"  
 /translation="MGAMGAAPLHSLVWKRRCVSLRPARALLRWRSPPRGP  
 PGDARSVLSEIIITAVEKDDCEKHASSGRWHKXENPPATVVRVKRVRHVRWKWAR  
 TFWASDQLCHLMQTLRGLLESLSRPKLLVFINPFGGQGGQKRIYKIVAPLFTL  
 ASITEIIITTEHANQAKETIYEINTSDYGVCGDGMFSEVLHGVIQTQSGAD  
 PNHPRAVLVSTLRIGIIPAGSTDVCYSTVGTNDASTALHIIIGDSLADVSSVHY  
 HNTLLRYSVLLGYFGDILIKDSEKKRMGLVRYDPSGLKTFLSHOYREGTLSFLPA  
 QHTVGSPRDNKPCRAGFCVCRSQKOLEEKKALYLENAEEVEEMOVTCGKFLA IN  
 ATNMSCAPRSPGSLSPAHLGQSSDLILIRKSRNFLRIHRTNQEDQDFTFV  
 EVRVKPHFTSKHVEDNDSRQEKQKFGKICKDRPSTCSASRSSMNCNDEGVMS  
 PAIEVRVHCQLVRLFARGIBEEB"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 8.32e-191 Length: 2830  
 Score: 2436.50 Matches: 454  
 Percent Similarity: 90.19% Conservative: 33  
 Best Local Similarity: 84.07% Mismatches: 48  
 Query Match: 84.37% Indels: 5  
 DB: 10 Gaps: 1

US-10-631-958-10 (1-537) x AB079067 (1-2830)

Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnGlnArg 20  
 Db 1 ATGGGGCAATGGGGCGGCGAGCGCTGCACCTCCGTGCTGGTGAACGCGCAGC 60  
 Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40  
 Db 61 TGTCCGCTCAGCTGGAGCCGCGCGGGCTCTGCTACGCTGGTGGCGAGCCGAGCCC 120  
 Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluLeuAlaVal 60  
 Db 121 GGGCCCTCCGCGCGGGTGGCGATGCCGCTCGGTACTGGTGTGGAGATCATCGCTGT 180  
 Qy 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80  
 Db 181 GAGGAAAAAGACGACTGCGAGAAACACAGCCCTCCAGTGGCGGATGCATATAAATGGAAT 240  
 Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAla 100  
 Db 241 CCGTTCGCATTCACAGCTCCACCGTGTGAAGCGAGTACGACATCACCGCTGGAAGTGGCA 300  
 Qy 101 GlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120  
 Db 301 CGGTGTGACCTTCTGGAGCGCGAGCAGCAGCTGTGTACCTGTGGCTTGCAGACCTCCGT 360  
 Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheLeuAsnProPhe 140  
 Db 361 GGGCTGTGGAGAGCCTGACTTCAAGACCGAAGCACTTGTCTGGTGTATTCATCAACCCCTTC 420  
 Qy 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160  
 Db 421 GGAGGGAAGGTTCAGGGCAAGCGCATCTATGAAAAAACAAGTGGCGCCTCTGTATTACCTTG 480

161 AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180  
 Db 481 GCTTCCATCACTACGGAGATCATATTACAGACATGCCAACCAAGCAAGGACATTTA 540  
 Qy 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200  
 Db 541 TACGAGATCAACACACAGACAGCTATGATGGCATCGTGTGCTGTGGTGGGACGGCATGTT 600  
 Qy 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220  
 Db 601 AGCGAGGTGCTGATGGGGTGATTGGAGAGCAGCAGAGAGCGCTGGTATTCGACCCCAAT 660  
 Qy 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240  
 Db 661 CACCCCGAGCGCTGCTGGTGGCCAGTACCCTCAGGATCGGCATCATACC CGCAGGGTCC 720  
 Qy 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260  
 Db 721 ACAGATTGTGTGTTACTCAACAGTAGGCACAACAGCAGCAGACATCGGCTTTGCAT 780  
 Qy 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 280  
 Db 781 ATCATTATTGGGACCTACCTGGCAATAGACGTGTCTCTGTGCATACCATTAACACCGCTG 840  
 Qy 281 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer 300  
 Db 841 CTGGGTACTCGGTTCTCTGCTGGGTACGGTTCCTACGGGACATTAAATCAAGGACAGC 900  
 Qy 301 GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 320  
 Db 901 GAAAAAGAACCGTGGATGGGCTCGTCCGGTATGATTTCTCAGGGTTGAAGACCTTTCTC 960  
 Qy 321 SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340  
 Db 961 TCTCATCAGTACTATGAAGGACACTGTCTCTCTCCAGCAGACACACGGTGGGATCT 1020  
 Qy 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360  
 Db 1021 CCACGGGACAATAAACCTGCCGGCTGGGTGCTCTGTGTGGCAGCAGCAAGCAACAG 1080  
 Qy 361 LeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 380  
 Db 1081 CTGGAAGAAGAAGAAGAACCCCTGTATGGCTGGAGAACCGCCGGAAGTGGGAAG 1140  
 Qy 381 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400  
 Db 1141 TGGCAAGTGACATGTGGAAAGTTCCTGGCCATCAATGCCACCAACATGTCTGTGTGT 1200  
 Qy 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420  
 Db 1201 CCTCGGAGCCCTGGGGCCCTGCCCATTTGCCATCTGGGAGATGGGTCTCTTGACCTC 1260  
 Qy 421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440  
 Db 1261 ATCCTTATCCGGAAGTGTCCAGGTTCACCTTCAGATTTCTCATCCGSCACACGAAC 1320  
 Qy 441 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 460  
 Db 1321 CAGAGAGCAGTTCGACTTCACCTTCGTTGAAGTTATTCGAGTCAAGAAATTCACCTTC 1380  
 Qy 461 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe 480  
 Db 1381 ACCTCGAGACCGCTGGAGATGAGGACATGACTCGAAGGAACAAGAGAGCAAGATTT 1440  
 Qy 481 GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn 500  
 Db 1441 GGGAAAGATCTCGAAGGACAGACCCCTCTTGCACTTGCTCAGCCTCCAGAACTCTCTGGAAC 1500  
 Qy 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520  
 Db 1501 TGCATGGCGAAGTCAATGACACGCCCGGCCATTTGAGGTGAGGGTCCATCTCCAGCTGGTG 1560



```

Qy 521 ArgLeuPheAlaArgGlyIleGluGlu-----AsnProIysProAspSer 535
Db 1561 CGCCTCTTTCGTCGGGAATCAGGAGAGTCATAGCAAGAACCCCAAGCCCGAGGAGC 1620

RESULT 13
AKI29416 3661 bp mRNA linear ROD 21-NOV-2003
LOCUS Mus musculus mRNA for mKIAA1646 protein.
DEFINITION AKI29416
ACCESSION AKI29416.1 GI:37360495
VERSION FLI CDNA.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus

REFERENCE
AUTHORS Okazaki,N., Kikuno,R., Ohara,T., Inamoto,S., Koseki,H., Hiraoka,S.,
Saga,Y., Nagase,T., Ohara,O. and Koga,H.
TITLE Prediction of the coding sequences of mouse homologues of KIAA
gene: iii. the complete nucleotide sequences of 500 mouse
KIAA-homologous cDNAs identified by screening of terminal sequences
of cDNA clones randomly sampled from size-fractionated libraries
DNA Res. 10 (4), 167-180 (2003)
JOURNAL 22977043
MEDLINE 14621295
PUBMED
REFERENCE 2 (bases 1 to 3661)
AUTHORS Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.
TITLE Direct Submition
JOURNAL Submitted (23-JUL-2003) Hisashi Koga, Kazusa DNA Research
Institute, Laboratory for Genome Informatics; 2-6-7
Kazusa-kanatari, Kisarazu, Chiba 292-0818, Japan
(E-mail:mous@kazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)
COMMENT The CREATE program supported by Japan science and technology
corporation; cDNA full insert sequencing: Kazusa DNA Research
Institute; cDNA library construction, clone selection and 5'- &
3'-end one pass sequencing.
FEATURES
source
location/Qualifiers
1..3661
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="mph01246"
/tissue_type="embryonic tail"
/note="Vector:modified pBC SK+"
1..3661
/gene="mKIAA1646"
<1..1231
/gene="mKIAA1646"
/note="CDS is predicted by in silico analysis. Start codon
is not identified.
Start codon is not identified."
/codon_start=2
/evidence=not experimental
/product="mKIAA1646 protein"
/protein_id="BAC98226.1"
/db_xref="GI:37360496"
IITTEHANQAKETIYEINTSDYGVICVCGDGMFSEVLHVGIRTOQSAGIDPNHRA
VLVPTLRIGIIPAGSTDVCYSTVGTNDASTALHIIIGDSLAIIDVSSVHYHNTLLR
YSVLGCGFYGDLLKDKSEKRWGLVRYDPSGLKFLSHQYVSGTSLFPLPAQHTVGS
PRDNKPCRGCFVCRKQLEBEKKALYGLNAAVEEWMQVTCGKFLAINATMNC
ACPRSPGLSPFALHGDGSSDLILIRKCSRFNLFRLIRHMQDQDFTFVEVYRYK
KPFHTSHVEDENDSKQEKKQFKGI CKDRPSCSTCSASRSSMNCDEGMHSPAIEVR
VHCQLVRLPARGIBES"

ORIGIN
Alignment Scores:
Pred. No.: 1.81e-147 Length: 3661
Score: 1912.50 Matches: 356
Percent Similarity: 92.49% Conservative: 26
Best Local Similarity: 86.20% Mismatches: 26
Query Match: 66.22% Indels: 5
DB: 10 Gaps: 1

```

```

US-10-631-958-10 (1-537) x AKI29416 (1-3661)
Qy 128 SerArgProIysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLys 147
Db 17 TCAGAGCCGAAGCACTTGTGCTATTATCATCACCCCTTCGAGGGGAAAGGTCCGGGCAAG 76
Qy 148 ArgIleTyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIle 167
Db 77 CGCATCTATGAAAAACAGTGGCGCCTCTGTTTACCTTGGCTTCATCCTACGAGATC 136
Qy 168 IleValThrGluHisAlaAsnGlnAlaValGluThrLeuTyrGluIleAsnIleAspLys 187
Db 137 ATCAATTACAGAGCATGCCAACCAAGCAGAGAGACTTTATACGAGATCAACACAGACAGC 196
Qy 188 TyrAspGlyIleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeu 207
Db 197 TATGATGGCATCGTGTGCTAGGTGGGAGCGGCATGTTCCAGCGAGGTGCTGCATGGGGTG 256
Qy 208 IleGlyArgThrGlnArgSerAlaGlyValAspGlnAenHisProArgAlaValLeuVal 227
Db 257 ATTGGGAGGAGCGCAGCAGAGCGCTGGTATCGACCCCAATCACCCCGAGCGCTGCTGTG 316
Qy 228 ProSerSerLeuArgIleGlyIleProAlaGlySerThrAspCysValCysTyrSer 247
Db 317 CCCAGTACCCCTCAGGATCGGCATCATCCCGCAGGGTCCACAGATTGTGTGTACTCA 376
Qy 248 ThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeu 267
Db 377 ACAGTGGGCACNAACGACGACGACATCGGCTTTCATCATCATATTGGGGACTCGCTG 436
Qy 268 AlaMetAspValSerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeu 287
Db 437 GCATAGACGTGTCCTCTGTGCACTACATAAACCGCTGCTCGGTACTCGGTTCTCTG 496
Qy 288 LeuGlyTyrGlyPheTyrGlyAspIleIleIysAspSerGluLysLysArgTyrLeuGly 307
Db 497 CTGGGCTACGGTTTCTACGGGGACTTAATCAAGGACAGCGAAGAAACGGTGGATGGGC 556
Qy 308 LeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisCysTyrGluGly 327
Db 557 CTGTCGGGTATGATTTCTCAGGGTTGAGACCTTTCTCTCATCATGACTACTATGAGGG 616
Qy 328 ThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysProCys 347
Db 617 ACCTGTCTCTCCCGACAGACACACGCTGGGATCTCCACGGGACAAATAAACCCCTGC 676
Qy 348 ArgAlaGlyCysPheValCysArgGlnSerIysGlnLeuGluGluGlnLysLys 367
Db 677 CGGGCTGGGTGCTTCGTGTCAGGACAGCAAGCAACAGCTGGGAAGAGAGAGAGAAA 736
Qy 368 AlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGluTrpGlnValValCysGlyLys 387
Db 737 GCCCTGTATGGCTGGAGAACGCCGAGGAAGTGGAGAGTGGCAAGTGACATGTGGGAAG 796
Qy 388 PheLeuAlaIleAenAlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeu 407
Db 797 TTCCTGGCCATCAATGCCACCAACATGCTCTGTCTGTCTCTCGAGCCCTGGGGGCGCTG 856
Qy 408 SerProAlaAlaHisLeuGlyAspGlySerSerAspIleIleLeuIleArgLysCysSer 427
Db 857 TCCCATTTTCCCATCTCGGAGAGATGGGTCTTCTGACCTCATCTCTATCCGGAAGTGTCTC 916
Qy 428 ArgPheAsnPheLeuArgPheLeuIleArgHisThrAenGlnGlnAspGlnPheAspPhe 447
Db 917 AGTTTCAACTTCTGAGATTTCTCATCCGCCACACGAAACGAGGAGCCAGTTCGACTTC 976
Qy 448 ThrPheValGluValTyrArgValIysPheGlnPheThrSerLysHisMetGluAsp 467
Db 977 ACTTTCGTGAAGTTTATCGAGTCAAGAAATTCACATTCACGTCGAAGACGTCGAAGAC 1036
Qy 468 GluAspSerAspLeuLysGluGlyLysLysArgPheGlyHisIleCysSerSerHis 487

```





```

Db      1191  ATGAAGCAGTGTACACACATAGGAAGAAACATTTTGGCCAGTTCTGCAGATCAT 1250
Qy      488  ProSerCysCysThrValSerAenSerTrpAenCysApsGlyGluValLeuHis 507
Db      1251  CCAGCCTGCTGCTGAATATTGCTAATAGCACTTGGAAATTGTGTGGAGAACCTGGAT 1310
Qy      508  SerProAlaIleGluValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIle 527
Db      1311  AGCTCAGCAATTGAAGTGAAGGTTCCTGCCCTGTTAATGAACACTGTTTGCRAAGGGGAATC 1370

```

```

Qy      528  GluGluAenProLyProAap-----SerHisSer 537

```

```

Db      1371  GAGGAACACTGAAGATGAACAGTGCCTACACCCACAGT 1409

```

```

RESULT 15
BC074350
LOCUS      BC074350      2494 bp      mRNA      linear      VRT 03-AUG-2004
DEFINITION      Xenopus laevis MGC84197 protein, mRNA (cDNA clone MGC:84197
IMAGE:6954181), complete cds.
ACCESSION      BC074350
VERSION      BC074350.1      GI:49257664
KEYWORDS      MGC.
SOURCE      Xenopus laevis (African clawed frog)
ORGANISM      Xenopus laevis

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus; Xenopus.
1 (bases 1 to 2494)
Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W.
and Richardson,P.
Genetic and genomic tools for Xenopus research: The NIH Xenopus
Initiative
Dev. Dyn. 225 (4), 384-391 (2002)
12454917
2 (bases 1 to 2494)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marudina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schettz,T.E., Brownstein,M.J., Uesdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullighy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Pahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywicki,M.I., Skalek,U., Smalilus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
3 (bases 1 to 2494)
Klein,S. and Gerhard,D.S.
Direct Submission
Submitted (23-JUN-2004) National Institutes of Health, Xenopus Gene
Collection (XGC), National Institute of Child Health and Human
Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
20892-7510, USA
NIH-MGC Project
Contact: XGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement:
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada

```

```

info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabhu,
Parvaneh Saedi, JR Santos, Angeliq Schnerch, Ursula Skalska,
Duane Smalilus, Jeff Stott, Miranda Tsai, George Yang, Jacque
Schein, Asim Siddiqui, Rob Holt, Marco Marra.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: PRK Plate: 161 Row: a Column: 22  
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

#### FEATURES

```

source
1. 2494
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="MGC:84197 IMAGE:6954181"
/tissue_type="Brain, adult Xenopus"
/clone_lib="NICHD_XGC_Brn1"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6.1"
1. 2494
/gene="MGC84197"
/db_xref="LocusID:444656"
86..1804
/gene="MGC84197"
/codon_start=1
/product="MGC84197 protein"
/protein_id="AAH74350.1"
/db_xref="GI:49257665"
/db_xref="LocusID:444656"
/translation="MESAHTSLAVLCSVLVKQSCSEVTLDPGRSLLSWRELPRRH
KERSPGICLTVLHRTGLVNRRTFPFGSGVTVPVTEIVSGEAEIDEXYNSMKW
QHCRCPAFTVYVQRVGRKRCVETFWCSDQFLFYQWLQAFHDLLEQQTHRPKNL
LVYINPGGKRGKQIYENKQVPLFSAAGICADVITEYANHARDHLYDLNLEYKDV
VCGDGMFSEVLHGLVIRMQKSDVDHNNPSAOLSRNCNMRIIIPAGSTDCICVATV
GINDPETSALHILGDCQPLDVCSPVAKVFLKYSVLGVGYGVGLVKGSKNRWL
PARDVSGFTFLTHCYEGSVSPQAKWLGSPPRDQTTCTSGCYICRQSKQDDEQE
QYQVSGSEHREQDDTTIIGRFMAINAVMSACRSPNGSLSPAHLADGASDLIV
KRSCLDFRLHRLRHTSNKQDFPFVEYVRVKNFQTPKHFEDENESTDGLGKK
NFSQICTDHPSCGCHVNSIWCNCGETLDTAIEMRVHCQLIKLFARGIENAKLEDDV
NPSKVF"

```

#### gene

```

1. 2494
/gene="MGC84197"
/db_xref="LocusID:444656"
86..1804
/gene="MGC84197"

```

#### CDS

```

/codon_start=1
/product="MGC84197 protein"
/protein_id="AAH74350.1"
/db_xref="GI:49257665"
/db_xref="LocusID:444656"
/translation="MESAHTSLAVLCSVLVKQSCSEVTLDPGRSLLSWRELPRRH
KERSPGICLTVLHRTGLVNRRTFPFGSGVTVPVTEIVSGEAEIDEXYNSMKW
QHCRCPAFTVYVQRVGRKRCVETFWCSDQFLFYQWLQAFHDLLEQQTHRPKNL
LVYINPGGKRGKQIYENKQVPLFSAAGICADVITEYANHARDHLYDLNLEYKDV
VCGDGMFSEVLHGLVIRMQKSDVDHNNPSAOLSRNCNMRIIIPAGSTDCICVATV
GINDPETSALHILGDCQPLDVCSPVAKVFLKYSVLGVGYGVGLVKGSKNRWL
PARDVSGFTFLTHCYEGSVSPQAKWLGSPPRDQTTCTSGCYICRQSKQDDEQE
QYQVSGSEHREQDDTTIIGRFMAINAVMSACRSPNGSLSPAHLADGASDLIV
KRSCLDFRLHRLRHTSNKQDFPFVEYVRVKNFQTPKHFEDENESTDGLGKK
NFSQICTDHPSCGCHVNSIWCNCGETLDTAIEMRVHCQLIKLFARGIENAKLEDDV
NPSKVF"

```

#### ORIGIN

```

Alignment Scores:
Pred. No.: 3.08e-131      Length: 2494
Score: 1713.00      Matches: 320
Percent Similarity: 73.18%      Conservative: 81
Best Local Similarity: 58.39%      Mismatches: 117
Query Match: 59.31%      Indels: 30
DB: 5      Gaps: 4

US-10-631-958-10 (1-537) x BC074350 (1-2494)

Qy      10  LeuGlnSerValLeuTrpValLysGlnGlnArgCysAlaValSerLeuGluProAlaArg 29
Db      119  CTGTGCTCGGTGCTCTCGGTTAAGAGGCAAGCTGCGAGGTGACCTTGACCCCGGGCGC 178
Qy      30  AlaLeuLeuArgTrpTrpArgSerProGlyAlaGly-----AlaPro 45
Db      179  TCTCTGCTCTCC---TGGAGAGAACTGAGGCCCGCCACACGAGAGACGAGCGCGGCC 235
Qy      46  Gly-----
Db      236  GGAATCTGCTTGCAAGAGTGCTCCACCAACAAAGAGCGCTAGTAAACCGTTTCACTCGG 295
Qy      47  -----AlaAspAlaCysSerValProValSerGluIlelleAlaValGluGluThr 63

```



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 6, 2005, 02:19:54 ; Search time 829.053 Seconds  
(without alignments)  
3834.378 Million cell updates/sec

Title: US-10-631-958-10

Perfect score: 2888

Sequence: 1 MGATCAEPLQSLVWVKQOR.....QLVRLFARGIENKPDPSHS 537

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO spool/US10631958/runat 02092005 165814 4497/app query.fasta\_1.1941  
-DB=N Geneseq 16Dec04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPTCL=0  
-LOOPTXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 16Dec04: \*  
1: Geneseqn1980s: \*  
2: Geneseqn1990s: \*  
3: Geneseqn2000s: \*  
4: Geneseqn2001as: \*  
5: Geneseqn2001bs: \*  
6: Geneseqn2002as: \*  
7: Geneseqn2002bs: \*  
8: Geneseqn2003as: \*  
9: Geneseqn2003bs: \*  
10: Geneseqn2003cs: \*  
11: Geneseqn2003ds: \*  
12: Geneseqn2004as: \*  
13: Geneseqn2004bs: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2888	100.0	1740	8 ADA05679	Ada05679 Human NOV
2	2888	100.0	4413	6 ABL40828	Ab140828 Human sph
3	2888	100.0	4429	12 ADJ96598	Aqj96598 Human lip
4	2888	100.0	4445	13 ADP55247	Adp55247 Human PRO
5	2881	99.8	4432	8 ABX70921	Abx70921 Novel hum

6	2880	99.7	4463	6 ABA96945	Aba96945 Human cer
7	2843	98.4	1740	12 ADN62844	Adn62844 Human NOV
8	2649.5	91.7	1840	5 Aad14426	Aad14426 Human sph
9	2587	89.6	4231	3 AAA50510	Aaa50510 Human sph
10	2491	86.3	4702	13 ADS10370	Ads10370 Human the
11	2349.5	81.4	2241	5 AAS77728	Aas77728 DNA encod
12	2210	76.5	3975	8 ACC70838	Acc70838 Human sph
13	1744.5	59.4	2186	5 AAS77730	Aas77730 DNA encod
14	1640.5	56.8	979	6 ABL40822	Ab140822 Human sph
15	1386.5	48.0	1570	5 AAS77731	Aas77731 DNA encod
16	1055	36.5	753	3 AAC76031	Aac76031 Human ORF
17	1032	35.7	817	6 ABQ99499	Abq99499 Human cod
18	1032	35.7	817	13 ADS11669	Ads11669 Human the
19	986	34.1	547	3 AAC76592	Aac76592 Human ORF
20	871	30.2	550	6 ABL40827	Ab140827 Human sph
21	811	28.1	474	6 ABL40823	Ab140823 Human sph
22	710	24.6	10337	4 ABL25704	Ab125704 Drosophil
23	679	23.5	426	5 AAS77727	Aas77727 DNA encod
24	675	23.4	2173	4 ABL25705	Ab125705 Drosophil
25	656	22.7	522	5 AAD14427	Aad14427 Partial r
26	626	21.7	411	5 AAS77729	Aas77729 DNA encod
27	584	20.2	1774	12 ADQ88890	Adq88890 Novel hum
28	541.5	18.8	15185	4 AAK65589	Aak65589 Human imm
29	525	18.2	329	6 ABL40824	Ab140824 Human sph
30	511.5	17.7	15181	4 AAK65588	Aak65588 Human imm
31	437.5	15.1	564	8 ABT23453	Abt23453 Immune-re
32	343.5	11.9	382	5 AAD14428	Aad14428 Partial m
33	321	11.1	2830	4 ABL02327	Ab102327 Drosophil
34	319	11.0	2609	9 ACF35860	Acf35860 D. melano
35	310.5	10.8	2020	4 ABL03297	Ab103297 Drosophil
36	310.5	10.8	2629	9 ACF35859	Acf35859 D. melano
37	310.5	10.8	4020	4 ABL03296	Ab103296 Drosophil
38	302	10.5	167	6 ABL40825	Ab140825 Human sph
39	284.5	9.9	572	13 ACN57851	Acn57851 Cotton gy
40	281	9.7	2422	5 AAS85331	Aas85331 DNA encod
41	278	9.6	1857	6 ABK90199	Abk90199 cDNA enco
42	278	9.6	2380	5 AAS14817	Aas14817 Human cDN
43	278	9.6	2380	6 ABL59533	Ab159533 Human sph
44	278	9.6	2380	10 ADE85298	Ade85298 Farneyyl
45	278	9.6	2380	10 ADE38416	Ade38416 Human pro

ALIGNMENTS

RESULT 1	
ADA05679	
ID	ADA05679 standard; cDNA; 1740 BP.
XX	
AC	ADA05679;
XX	
DT	06-NOV-2003 (first entry)
XX	
DE	Human NOV9a encoding cDNA SEQ ID NO:39.
XX	
KW	human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
KW	immunomodulator; cytostatic; nootropic; neuroprotective;
KW	antiparkinsonian; antilipemic; gene therapy; human disease;
KW	metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
KW	neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW	immune disorder; haematopoietic disorder; dyslipidaemia; gene; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	CDS 76..1689
FT	/*tag= a
FT	/product= "NOV9a"
XX	
FN	WO2003029424-A2.
XX	
PD	10-APR-2003.
XX	
PF	02-OCT-2002; 2002WO-US031373.





```
Qy 81 ProTyrAlaPheThrValHisCysValIysArgAlaArgHisArgTTPLeuTIPAla 100
Db 316 CCTTACGCTTTACAGTTCACTGTGTAAAGAGACACGACCGCACCGCTGGAAGTGGCG 375
Qy 101 GlnValThrPheTTPCysProGluGluGlnLeuCysHisLeuTTPLeuGlnThrLeuArg 120
Db 376 CAGGTGACTTTCTCGTGTCCAGAGAGCAGCTGTGTCACTTGTGGCTGCAGACCTCGG 435
Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
Db 436 GAGATGCTGGAGAGCTGACGTCCAGACCAAGCATTTACTGGTATTATCAACCCCGTTT 495
Qy 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
Db 496 GGAGGAAAGACCAAGCAGCGATATATGAAGAAAAGTGGCACCACTGTTACCTTA 555
Qy 161 AlaSerIleThrThrAspIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180
Db 556 GCCTCCATCACCACCTGACATCATCGTTACTGAACATGCTAATCAGGCCAAGGACACTG 615
Qy 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200
Db 616 TATGAGATTAAACATAGACAAATACGACGGCATCGTGTGTGCGCGGAGATGGTATGTC 675
Qy 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220
Db 676 AGCAGGTGCTGCACGGTCTGATTGGAGGACGACAGAGGCGCGGGTCCAGCAGAAC 735
Qy 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240
Db 736 CACCCCGGGGTGTGTGTCCTCCAGTAGCCTCCGGATTTGGAAATCAATCCCGCAGGGTCA 795
Qy 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
Db 796 ACGGACTCGGTGTATTCTCCACCGTGGGCACCGACGACGAGAAACCTCGGCGCTGCAT 855
Qy 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 280
Db 856 ATCGTGTGGGACTCGCTGGCCATGGATGTGCTCCTCAGTCCACACACAGCAGCACTC 915
Qy 281 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer 300
Db 916 CTTGCTACTCCGTGTCCCTCGGCTTACGGCTTCTACGGGGACATCATCAAGGACAGT 975
Qy 301 GluLysLysArgTTPLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 320
Db 976 GAGAAGAAACCGTGTGGTCTTGCCAGATACGACTTTTCAGGTTTAAAGACCTTCCTC 1035
Qy 321 SerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340
Db 1036 TCCACCACTGCTATGAAGGACAGTGTCTTCTCCTCCACACACACCGTGGATCT 1095
Qy 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360
Db 1096 CCAAGGATAGGAAGCCCTGCGGCGCAGAGTGTGTGTTGTCAGGCCAAAGCAAGCAGCAG 1155
Qy 361 LeuGluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 380
Db 1156 CTGGAGGAGGACGAGAAGAACGACTGTATGTTTGGAAAGCTGCGGAGGAGCTGGAGGAG 1215
Qy 381 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400
Db 1216 TGGCAAGTCTGTGTGGGAAGTTCTGGCCATCAATGCCACAAACATGCTCTGTGCTGT 1275
Qy 401 ArgArgSerProArgGlyLeuSerProAlaIleHisLeuGlyAspGlySerSerAspLeu 420
Db 1276 CGCCGAGCCCGCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCTGACCTC 1335
Qy 421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440
Db 1336 ATCCTCATCCCGAAATGCTCCAGGTTCAAATTTTCTGAGATTCTCATCAGGCACCAAC 1395
```

```
Qy 441 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 460
Db 1396 CAGCAGGACCAGTTTGACTTCACTTTGTTGTAAGTTTATCGCGTCAAGAAATTCACGTTT 1455
Qy 461 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyGlyLysLysArgPhe 480
Db 1456 ACGTCTGAAGCACATGGAGGATGAGCAGCAGCGACTCTAAGAGGGGGGGAAGCGGCTTT 1515
Qy 481 GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTIPAsn 500
Db 1516 GGGCACATTTGCAGACCCACCTCTCTGCTGCTGCACCTCTCCAACAGCTCTCTGGAAC 1575
Qy 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520
Db 1576 TGGCAGCGGGAGGTCTCTGCACAGCCCTCGCCATCGAGGTCAAGTCCACTCCAGCTGGT 1635
Qy 521 ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 537
Db 1636 CGACTCTTTGCAGGAGGAAATTGAAGAGAATCCGAAGCCAGACTCACACAGC 1686

RESULT 3
ID ADJ96598 standard; DNA; 4429 BP.
XX AC ADJ96598;
XX XX 06-MAY-2004 (first entry)
XX DE Human lipid kinase KIAA1646 DNA SeqID 55.
XX KW gene; ds; kinase; human; SNP; single nucleotide polymorphism;
XX KW tyrosine protein kinase; serine/threonine protein kinase; PTK; STK;
XX KW gene therapy; cancer; immune-related disease; cardiovascular disease;
XX KW brain; neuronal associated disease; metabolic; inflammatory disorder;
XX KW cytosolic; neuroprotective; immunomodulator; antiinflammatory;
XX KW lipid kinase; KIAA1646.
OS Homo sapiens.
OS 39.
XX FH Location/Qualifiers
FT variation /*tag= a
FT /*standard_name= "Single nucleotide polymorphism"
FT variation /*tag= b
FT /*standard_name= "Single nucleotide polymorphism"
FT variation /*tag= c
FT /*standard_name= "Single nucleotide polymorphism"
FT variation /*tag= d
FT /*standard_name= "Single nucleotide polymorphism"
FT variation /*tag= e
FT /*standard_name= "Single nucleotide polymorphism"
XX WO2004006838-A2.
XX XX 22-JAN-2004.
XX PF 15-JUL-2003; 2003WO-US021730.
XX PR 15-JUL-2002; 2002US-0395632P.
XX FA (SUG- ) SUGEN INC.
XX PI Whyte D, Manning G, Caenepeel S;
XX DR WPI; 2004-122753/12.
XX DR P-PSDB; ADJ96664.
XX FT New nucleic acid molecule encoding a kinase polypeptide, useful for
```



PT preparing a composition for treating diseases or disorders, e.g., cancer,  
 XX or neurological, immunological or inflammatory disorders.

Example 1; SEQ ID NO 55; 366pp; English.

CC This invention relates to a novel isolated, enriched or purified nucleic  
 CC acid molecule that encodes a kinase polypeptide. Specifically, it relates  
 CC to human tyrosine and serine/threonine protein kinases (PK's and SRK's),  
 CC as well as protein kinase-like enzymes. The present invention describes  
 CC screening methods to identify agonists, antagonists and antibodies that  
 CC can be used to modulate the activity or function of the mammalian kinase  
 CC enzymes. As such, these compositions can be used for gene therapy  
 CC purposes to treat diseases or disorders including cancer, immune-related  
 CC diseases, cardiovascular disease, brain or neuronal associated disease,  
 CC metabolic and inflammatory disorders. Accordingly, they exhibit  
 CC cytostatic, neuroprotective, immunomodulator and antiinflammatory  
 CC activities. This polynucleotide sequence is a human kinase DNA sequence  
 CC of the invention.

XX  
 SQ Sequence 4429 BP; 1020 A; 1115 C; 1213 G; 1081 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.: 2,61e-254 Length: 4429  
 Score: 2888.00 Matches: 537  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0

US-10-631-958-10 (1-537) x ADJ96598 (1-4429)

QY 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValValGlnGlnArg 20  
 DB 92 ATGGGGGGACGGGGGGCGGAGCGCGTCAATCGTCTGTGGTGAAGCAGCGCGC 151  
 QY 21 CysAlaValSerLeuGluProAlaAlaLeuLeuArgTrpTrpArgSerProGlyPro 40  
 DB 152 TGCGCCGTGAGCTGGAGCCCGCGGGCTCTGCTGCGCTGGTGGCGAGCCCGGGGCC 211  
 QY 41 GlyAlaGlyAlaProGlyAlaAlaAspAlaCysSerValProValSerGluIleAlaVal 60  
 DB 212 GGAGCGGGCGCCCGCGCGGATCGCTCTGTGCTGTATCTGAGATCATCGCCGT 271  
 QY 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80  
 DB 272 GAGGAAACAGACGTTTCAAGGAAACATCAAGCAGTGGAAATGCGAGAAATGGAAG 331  
 QY 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrpAla 100  
 DB 332 CCTTACGCTTTTACAGTTTCACTGTGTAAGAGAGACACCGCACCGCTGGAAGTGGCG 391  
 QY 101 GlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120  
 DB 392 CAGGTGACTTCTGTGTCCAGAGAGCAGCTGTGTCTACTGTGCTGCAGACCCCTGCGG 451  
 QY 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140  
 DB 452 GAGATGCTGGAGAAGCTGACGCTCCAGACCAAGCATTTTACTGGTATTATCAACCCGTT 511  
 QY 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160  
 DB 512 GGAGGAAAGGAAAGGCAAGCGGATATATGAAGAAAGTGGCACCACTGTTCACCTTA 571  
 QY 161 AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180  
 DB 572 GCCTTCATCACCACCTGACATCATCTGTTACTGAACTGTCTATCAGGCCAAGGAGACTCTG 631  
 QY 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyLysGlyMetPhe 200  
 DB 632 TATGAGATTTAATAGACAAATACACGCGCATCGTCTGTGTCGGCGGAGATGTATGTC 691  
 QY 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220

DB 692 AGCAGGTTGTCGACGGTCTGATTGGGAGGAGCAGCAGGAGCGCGGGGTGCACCAGAAC 751  
 QY 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAlaGlySer 240  
 DB 752 CACCCCGGGCTGTGCTGGTCCCAGTAGCTCCCGATTGGAATCATTTCCCGAGGTCAC 811  
 QY 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260  
 DB 812 ACGACTCGGTGTGTACTCCACCGTGGCACCAGCAGCAGCAAACTCCGCGCTGCAT 871  
 QY 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 280  
 DB 872 ATCGTTGTTGGGACTCGCTGGCCATGGATGTCTCTCAGTCCACCAACACAGCACATC 931  
 QY 281 LeuArgTyrSerValSerLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer 300  
 DB 932 CTTTCGCTACTCCGTGCTCCTGCTGGCTACGGCTTACGGGGACATCATCAAGGACAGT 991  
 QY 301 GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 320  
 DB 992 GAGAAGAAACCGTGTGGTCTTCCACAGATACGACTTTTTCAGGTTTAAAGACCTTCCTC 1051  
 QY 321 SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340  
 DB 1052 TCCACCACTCTATGAGGAGACAGTGTCTTCTCTCCCTGCACCAACACACCGTGGATCT 1111  
 QY 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360  
 DB 1112 CCAAGGATAGGAAGCCCTGCGGGCAGGAGTCTTGTGTCAGGCAAGCAAGCAGCAG 1171  
 QY 361 LeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 380  
 DB 1172 CTGGAGGAGGAGCAGAAAGAACACTGTATGTGTTGGAAAGCTGCGAGAGAGCTGGAGGAG 1231  
 QY 381 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400  
 DB 1232 TGGCAAGTCGTCTGTGGAAAGTTTCTGGCCATCAATGCCAACAAATCTCTGTGCTGT 1291  
 QY 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420  
 DB 1292 CGCCGAGCCCGCAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCTGACCTC 1351  
 QY 421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440  
 DB 1352 ATCCTCATCCGGAATGCTCCAGGTTCAAATTTTCTGAGATTCTCATCAGGCACACCAAC 1411  
 QY 441 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 460  
 DB 1412 CAGCAGGACCAAGTTTGACTTCACCTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTT 1471  
 QY 461 ThrSerLysHisMetGluAspSerAspLeuLysGluGlyLysLysArgPhe 480  
 DB 1472 ACCTCGAAGCACATGGAGGATGAGACAGCAGCCTCAAGGAGGGGGGGAAGAGCGCTTT 1531  
 QY 481 GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn 500  
 DB 1532 GGGCACATTTTCAGCAGCAGCCACCTCTCTGCTGCTGCACCGTCTCCACACAGCTCTCTGGAAC 1591  
 QY 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520  
 DB 1592 TGCAGCGGGAGGTCTGTCACAGCCCTGCCATCAGAGGTTCAGAGTCCACTGCCAGCTGGTT 1651  
 QY 521 ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 537  
 DB 1652 CGACTCTTTGACCGAGGAAATTGAAGAGAAATCCGAAGCCAGACTCACAGC 1702

#### RESULT 4

ADP55247  
 ID ADP55247 standard; cDNA; 4445 BP.  
 XX  
 AC ADP55247;  
 XX  
 DT 18-NOV-2004 (first entry)

XX Human PRO cDNA sequence SEQ ID NO:1223.

DE human; PRO; immune related disease; inflammatory immune response;

XX immune response stimulation; antiallergic; antianemic; antiarthritic;

KW antiasthmatic; antidiabetic; antinflammatory; antiporiatic;

KW antirheumatic; antithyroid; CNS; dermatological; gastrointestinal;

KW haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;

KW nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;

KW virucide; gene therapy; gene; ss.

XX

OS Homo sapiens.

XX WO2004039956-A2.

PN 13-MAY-2004.

PD 28-OCT-2003; 2003WO-US034381.

PF 29-OCT-2002; 2002US-0422472P.

XX (GETH ) GENENTECH INC.

XX Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;

PI Wood WI, Wu TD;

XX WPI; 2004-376182/35.

DR P-PSDB; ADP55248.

XX

XX New PRO polynucleotides and polypeptides, useful in diagnosing

PT and treating an immune related disease, e.g. systemic lupus

PT erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in

PT stimulating an immune response.

XX

PS Claim 2; SEQ ID NO 1223; 3009pp; English.

XX

CC The present invention describes an isolated PRO nucleic acid (1). Also

CC described: (1) a vector comprising (1); (2) a host cell comprising the

CC vector of (1); (3) a process for producing a PRO polypeptides; (4) an

CC isolated PRO polypeptide; (5) a chimeric molecule comprising the

CC polypeptide of (4) fused to a heterologous amino acid sequence; (6) an

CC antibody which specifically binds to a polypeptide of (4); (7) a

CC composition of matter comprising a polypeptide of (4), an agonist or

CC antagonist of the polypeptide or an antibody that binds to the

CC polypeptide in combination with a carrier; (8) an article of manufacture

CC comprising a container, a label on the container and a composition of

CC matter of (7); (9) a method of treating an immune related disease in a

CC mammal; (10) a method for determining the presence of a PRO polypeptide

CC in a sample suspected of having the polypeptide; (11) a method of

CC diagnosing an immune related disease or an inflammatory immune response

CC in mammal; (12) a method of identifying a compound that inhibits or

CC mimics the activity of or expression of a gene encoding a PRO polypeptide

CC ; and (13) a method of stimulating the immune response in a mammal. The

CC PRO sequences have antiallergic, antianemic, antiarthritic,

CC antiasthmatic, antidiabetic, antinflammatory, antiporiatic,

CC antirheumatic, antithyroid, CNS, dermatological, gastrointestinal,

CC haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,

CC nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and

CC virucide activities, and can be used in gene therapy. The nucleic acid

CC (1) and the encoded polypeptides, compositions, kits and methods are

CC useful in diagnosing and treating an immune related disease and in

CC stimulating an immune response. The present sequence represents a human

CC PRO nucleotide sequence from the present invention.

XX

SQ Sequence 4445 BP; 1018 A; 1134 C; 1216 G; 1077 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.63e-254	Length:	4445
Score:	2888.00	Matches:	537
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	13	Gaps:	0

US-10-631-958-10 (1-537) x ADP55247 (1-4445)

Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnArg 20

Db 124 ATGGGGCGACGGGGCGGGAGCCGCTGCAATCCGTGCTGTGGTGAACAGCAGCGC 183

Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40

Db 184 TGGCGCGTGAGCCCTGGAGCCCGCGGGGCTCTGCTGCGCTGGTGGCGGAGCCCGGGGCC 243

Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluLeuLeuAlaVal 60

Db 244 GGAGCCGGCGCCCGCGCGGGATGCCCTGCTCTGTGCTGTATCTGAGATCATCGCGTT 303

Qy 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80

Db 304 GAGGAAACAGACCGTTACCGGGAACATCAAGGGCAGTGGAAATGGCAGAAATGGAAAG 363

Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrpLysTrpAla 100

Db 364 CCTTACGCTTTTACAGTTCACTGTGTAAAGAGACGACGCGCACCGCTGGAAGTGGCG 423

Qy 101 GlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120

Db 424 CAGGTGACTTTCTGCTGTCAGAGGAGCAGCTGTGTCACTTGTGCTGCAGACCTGCGG 483

Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheLeuAsnProPhe 140

Db 484 GAGATGCTGGAGAGCTGACGCTCCAGACCAAGCATTTACTGTGTTATTTATCAACCCGTT 543

Qy 141 GlyGlyLysGlyGlnGlyLysArgLysArgLysValAlaProLeuPheThrLeu 160

Db 544 GGAGGAAAGAGCAGCAGGCGGATATGAAAGAAAAGTGGCACCACCTTACCTTA 603

Qy 161 AlaSerIleThrAspIleIleValThrGluHisAlaAlaGlnAlaLysGluThrLeu 180

Db 604 GCCTCCATCACCACGACATCATCGTTACTGAACATGCTAATCAGGGCAAGAGACTCTG 663

Qy 181 TyrGluIleAsnIleAspLysTrpAspGlyLysValCysValGlyGlyAspGlyMetPhe 200

Db 664 TATGAGATTAACTAGACAAATACGAGCGCATCGTCTGTGTGGCGGAGATGTTATGTT 723

Qy 201 SerGluValLeuHisGlyLeuLeuGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220

Db 724 AGCGAGTGTGTCACGCTCTGATTGGGAGGACGACAGAGGCGCGGGGTGACACGAGAC 783

Qy 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyLysIleProAlaGlySer 240

Db 784 CACCCCGGGCTGTGTGTCCTCCAGTAGCTCCCGATTGGAATCATTTCCCGCAGGGTCA 843

Qy 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260

Db 844 ACGGACTGCGGTGTGTACTCCACGTCGGGACCCAGCGACGAGAAACCTCGGCGCTGAT 903

Qy 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 280

Db 904 ATCGTTGTTGGGACTCGCTGGCCATCGATGTGCTCTCAGTCCACACACACAGCAGACTC 963

Qy 281 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleLeuLysAspSer 300

Db 964 CTTGCTACTCCGTGCTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAGGACAGT 1023

Qy 301 GluLysIleArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 320

Db 1024 GAGAAAGAACGGTGGTGGGTCTTGGCCAGATACACATTTTCAGGTTTTAAAGACCTTCCTC 1083

Qy 321 SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340

Db 1084 TCCACCCACTGTCTATGAAGGGACAGTGTCTCTTCTCTCCCTGCACACACACGCTGGATCT 1143

Qy 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360

Db 1144 CCAAGGATAGGAGCCCTGCGGGCAGGATGCTTTGTTTCAGGCAAGCAAGCAGCAG 1203  
Qy 361 LeuGluGluGluGlnLysLysAlaLeuYrGlyLeuGluAlaAlaGluAspValGluGlu 380  
Db 1204 CTGGAGGAGGAGCAGAGAAAGCAGCTGTATGTTTGGAGCTGCGGAGGACGTGGAGGAG 1263  
Qy 381 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400  
Db 1264 TGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCCTGTCTGTGT 1323  
Qy 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyValSerSerAspLeu 420  
Db 1324 CGCGGAGCCCGAGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCTGACCTC 1383  
Qy 421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440  
Db 1384 ATCTCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTCTCATCAGGCACCAAC 1443  
Qy 441 GlnGlnAspGlnPheAspPheThrPheValGluValTyArgValLysLysPheGlnPhe 460  
Db 1444 CAGCAGGACCACTTTGACTTCACTTTTGTGTGAAGTTTATCGCGTCAAGAAATTTCCAGTTT 1503  
Qy 461 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe 480  
Db 1504 AGCTCGAGGCATCGAGATGAGCAGCAGCCTCAAGGAGGGGGGAAGAGCGCTTT 1563  
Qy 481 GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn 500  
Db 1564 GGGCACATTTGCGAGCAGCCACCCCTCTGCTGTCGACCGTCTCCACAGCTCCTTGGAAC 1623  
Qy 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520  
Db 1624 TCGCATGGGGAGGTCTGTCACAGCCCTGCCATCGAGGTGAGAGTCCACTGCCAGCTGGTT 1683  
Qy 521 ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 537  
Db 1684 CGACTCTTTGACGAGGAATTGAGAGATCCGAAGCCAGACTCACAGC 1734

RESULT 5

ID ABX70921 standard; cDNA; 4432 BP.  
XX  
AC ABX70921;  
XX  
DT 05-MAR-2003 (first entry)  
XX  
DE Novel human cDNA sequence #146.  
XX  
KW Human; gene; ss; nervous system disorder; peripheral neuropathy;  
KW Huntington's disease; amyotrophic lateral sclerosis; haemophilia;  
KW neurodegenerative disease; Parkinson's disease; Alzheimer's disease;  
KW autoimmune disease; systemic lupus erythematosus; rheumatoid arthritis;  
KW insulin-dependent diabetes mellitus; anaemia; thrombocytopaenia; wound;  
KW ulcer; burn; bone disorder; osteoporosis; osteoarthritis; stroke;  
KW fibrosis; reperfusion injury; infection; allergic rhinitis; asthma;  
KW coagulation disorder; cancer; tumour; inflammatory disease; septic shock;  
KW Crohn's disease; anaphylaxis; proliferation; chemotactic;  
KW differentiation; stem cell growth factor; haematopoiesis; chemokine;  
KW haemostatic; antiinflammatory; expressed sequence tag; EST.  
XX  
OS Homo sapiens.  
XX  
PN WO200281731-A2.  
XX  
PD 17-OCT-2002.  
XX  
PF 29-JAN-2002; 2002WO-US001222.  
XX  
PR 30-JAN-2001; 2001US-00774528.  
XX  
PA (HYSE-) HYSEO INC.  
PA (GOOD/) GOODRICH R W.  
XX

PI Tang TY, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
PI Xue AJ, Yang Y, Wehrman T, Wang J, Wang D, Drmanac RT;  
XX WPI; 2003-058563/05.  
XX  
PT Novel polypeptide useful for treating neurodegenerative diseases, myeloid  
PT or lymphoid cell disorders, bone disorders, mechanical and traumatic  
PT disorders, coagulation disorders, and inflammatory diseases.  
XX  
PS Claim 1; Page: 612pp; English.  
XX  
CC This invention relates to the cDNA sequences encoding an isolated novel  
CC human polypeptide. The protein encoded by the nucleic acid of the  
CC invention is useful for treating central and peripheral nervous system  
CC diseases (e.g. peripheral neuropathy, Huntington's disease, amyotrophic  
CC lateral sclerosis); neurodegenerative diseases (e.g. Parkinson's disease,  
CC Alzheimer's disease); autoimmune disease (e.g. systemic lupus  
CC erythematosus, rheumatoid arthritis, insulin-dependent diabetes mellitus)  
CC ; myeloid or lymphoid cell disorders (e.g. anaemia and thrombocytopaenia)  
CC ; wounds, ulcers, burns; bone disorders (e.g. osteoporosis,  
CC osteoarthritis); mechanical and traumatic disorders (e.g. stroke, head  
CC trauma); lung or liver fibrosis; reperfusion injury in various tissues;  
CC bacterial, viral or fungal infections; allergic conditions such as  
CC allergic rhinitis, asthma; coagulation disorders (e.g. haemophilia);  
CC cancer and tumours; and inflammatory diseases (e.g. Crohn's  
CC disease, anaphylaxis). The protein may be used to inhibit the growth,  
CC infection or function of infectious agents such as bacteria, fungi,  
CC viruses, or to effect bodily characteristics, biorhythms or circadian  
CC cycles of rhythms. The protein may also have  
CC proliferation/differentiation, stem cell growth factor, haematopoiesis  
CC regulation, immune stimulating or suppressing, chemotactic/chemokinetic,  
CC haemostatic and thrombolytic, receptor/ligand, and antiinflammatory  
CC activities. The cDNA sequences of the invention are useful for expressing  
CC recombinant protein for analysis. The present sequence represents a novel  
CC human cDNA sequence of the invention, this sequence is an expressed  
CC sequence tag (EST) and was identified using subtractive hybridisation  
XX  
SQ Sequence 4432 BP; 970 A; 1122 C; 1282 G; 1058 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.15e-253	Length:	4432
Score:	2881.00	Matches:	536
Percent Similarity:	99.81%	Conservative:	0
Best Local Similarity:	99.81%	Mismatches:	1
Query Match:	99.76%	Indels:	0
DB:	8	Gaps:	0

US-10-631-958-10 (1-537) x ABX70921 (1-4432)

Qy	1	MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnGlnArg	20
Db	1204	ATGGGGGCGACGGGGCGCGGAGCGCGTCAATCCGTGCTGGTGAAGCAGCAGCGC	1263
Qy	21	CysAlaValSerLeuGluProAlaArgAlaLeuArgTrpTrpArgSerProGlyPro	40
Db	1264	TGCGCCGTGAGCCCTGGAGCCCGCGCGGCTCTGCTGCGCTGCGGAGCGCGGGGCC	1323
Qy	41	GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal	60
Db	1324	GGAGCCGGGCGCCCGGGGGGAGTGCCTGCTGTGCTGTATCTGAGATCATCCCGTT	1383
Qy	61	GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys	80
Db	1384	GAGGAAACAGACGTTTCACGGGAAACATCAAGGCAGTGGAAAATGGCAAAAATG	1443
Qy	81	ProTyrrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAla	100
Db	1444	CCTTACGCTTTTACAGATTCTGTGTAAAGAGACGACGACCGCTGGAAGTGGCG	1503
Qy	101	GlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg	120
Db	1504	CAGGTGACTTTCTGTGTGTCCAGAGGAGCAGCTGTGTCTGTGTGTGTGTGTGTGT	1563

Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140  
Db 1564 GAGATGCTGGAGAAGCTGACGTCCAGACCAAGCATTTACTGGTATTATCAACCCGTTT 1623  
Qy 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160  
Db 1624 GGAGGAAAGACCAAGGCAAGCGGATATATGAAGAAGAGTGGCCACCACTGTTTCACTTA 1683  
Qy 161 AlaSerIleThrThrAspIleIleValThrGluHisAlaIleGlnAlaLysGluThrLeu 180  
Db 1684 GCCTCCATCACCATGACATCATCGTTACTGAAACATGCTAATCAGGCCAAGAGACTCTG 1743  
Qy 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200  
Db 1744 TATGAGATTACATAGACAAATACGCGGCATCGTCTGTGCGCGGAGATGGTATGTTC 1803  
Qy 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220  
Db 1804 AGCGAGGTGCTGCACGGTCTGATTGGGAGGACGAGAGGCGCGGGTCCAGCAGAAC 1863  
Qy 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240  
Db 1864 CACCCCGGGCTGTGCTGCTCCAGTAGCTCCGGATTGGAATCATTTCCCGCAGGGTCA 1923  
Qy 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260  
Db 1924 ACGNACTCGGTGTACTCCACCGTGGCACACGAGCGCAGAAACCTCGGCGCTGCAT 1983  
Qy 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisAsnSerThrLeu 280  
Db 1984 ATCGTTGTGGGACATCGCTGGCCATGATGTCTCAGTCCACCAACAGCAGCACTC 2043  
Qy 281 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer 300  
Db 2044 CTTGCGTACTCCGTGCTCGCTGCGCTACCGGGACATCATCAAGGAGAGT 2103  
Qy 301 GluLysLysArgTyrLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 320  
Db 2104 CAGAAGAACCGTGGTGGTCTGCCAGATACGACTTTTCAGGTTTAAAGACCTTCTC 2163  
Qy 321 SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340  
Db 2164 TCCACCACTGCTATGGAGGACAGTGTCTTCTCCTCGCACCAACACACCGGTGGGATCT 2223  
Qy 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360  
Db 2224 CCAAGGGATAGAAAGCCCTGCGGCGAGATGCTTTGTTTCAGCGCAAGCAACGACGAG 2283  
Qy 361 LeuGluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 380  
Db 2284 CTGGAGGAGGAGCAGAAAGCACTGTATGTTTGGAAAGTCTCGGAGGACGTGGAGGAG 2343  
Qy 381 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400  
Db 2344 TGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGCTCTGTGTGT 2403  
Qy 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420  
Db 2404 CGCGGAGCCCCAGGGGCTCTCCCGGCTGCCACATTTGGGAGACGGTCTCTGACCTC 2463  
Qy 421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440  
Db 2464 ATCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCACACCAAC 2523  
Qy 441 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 460  
Db 2524 CAGCAGGACCAAGTTTGAAGTTTCACTTTTGTGAAGTTTATCGCGTCAAGAAATTCACGTTT 2583  
Qy 461 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe 480  
Db 2584 ACGTCGAAGCATGGAGATGAGACAGACCTCAAGGAGGGGGGAAAGAGCGCTTT 2643  
Qy 481 GlyHisIleCysSerSerHisProSerCysCysCysThrValSerAsnSerSerTrpAsn 500

Db 2644 GGGCACATTTGCAGCAGCCACCCCTCTCTGCTGTGCACTCTCCACAGCTCTCGAAC 2703  
Qy 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520  
Db 2704 TGGACGGGAGGTCTTCGACAGCCCTGCCATCGAGTCCAGTCCAGCTGGTT 2763  
Qy 521 ArgLeuPheAlaArgGlyIleGluGluAsnProLysProLysProAspSerHisSer 537  
Db 2764 CGACTCTTGCACGAGGAATTGAGAGAAATCCGAGGCCAGACTCACACAGC 2814  
RESULT 6  
ABA96945  
ID ABA96945 standard; cDNA; 4463 BP.  
XX ABA96945;  
AC ABA96945;  
XX 20-MAY-2002 (first entry)  
DT Human ceramide kinase hCERK1-encoding cDNA.  
XX  
XX Homo sapiens.  
OS  
XX Key Location/Qualifiers  
FH 124..1737  
FT /\*tag= a  
FT /product= "Human ceramide kinase hCERK1"  
XX  
XX WO200196575-A1.  
XX 20-DEC-2001.  
XX 11-JUN-2001; 2001WO-JP004889.  
XX 14-JUN-2000; 2000JP-00178039.  
XX (SANY ) SANKYO CO LTD.  
XX Sugiura M, Kono K, Kohama T;  
XX WPI; 2002-179513/23.  
XX P-PSDB; AAM49115.  
XX Human ceramide kinase gene and the enzyme encoded by it for screening  
XX substances as drugs for neurological, inflammatory and other disorders.  
XX Claim 5; Page 46-53; 61pp; Japanese.  
XX This sequence represents cDNA encoding a human ceramide kinase designated  
XX hCERK1. The invention relates to hCERK1, nucleic acids encoding it,  
XX expression vectors and host cells containing hCERK1 nucleic acids, the  
XX recombinant production of hCERK1 and antibodies specific for hCERK1. The  
XX invention also encompasses methods of isolating hCERK1 from samples, the  
XX use of hCERK1 in drug screening, and the use of hCERK1 nucleic acid  
XX sequences in gene therapy. hCERK1 mediates the ATP-dependent 1-  
XX phosphorylation of ceramides and can be used to screen for therapeutic  
XX and preventive agents for a wide range of disorders. Such disorders  
XX include neurological disease, inflammation, human immunodeficiency virus  
XX (HIV) infection, type 2 diabetes, obesity, sepsis, arteriosclerosis and  
XX cancer  
SQ Sequence 4463 BP; 1026 A; 1135 C; 1217 G; 1084 T; 0 U; 1 Other;  
Alignment Scores: 1.43e-253 Length: 4463  
Pred. No.: 2880.00 Matches: 535  
Score:

Percent Similarity: 99.81%				Conservative: 1			
Best Local Similarity: 99.63%				Mismatches: 1			
Query Match: 99.72%				Indels: 0			
DB: 6				Gaps: 0			
US-10-631-958-10 (1-537) x ABA96945 (1-4463)							
Qy	1	MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrrpValLysGlnArg	20	Db	1084	TCCACCACTGCTATGAAGGACAGTGTCTCTCTCTCCCTGCACACACACCGTGGATCT	111
Db	124	ATGGGGGCGAGCGGGGCGGCGAGCCGCTGCAATCCGCTGTGGGTGAAGCAGCAGCGC	183	Qy	341	ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln	360
Qy	21	CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrrpTrpArgSerProGlyPro	40	Db	1144	CAAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTGCAGGCAAAAGCAAGCAGCG	120
Db	184	TGGCGCGTGAGCCTGGAGCCCGCGCGGCTCTGCTGCGCTGTGGCGAGCCCGGGGCC	243	Qy	361	LeuGluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu	380
Qy	41	GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleIleAlaVal	60	Db	1204	CTGGAGGAGGAGCAGAAAGAACCTGTATGGTTTGAAGCTGCGGAGGACGTGGAGGAG	126
Db	244	GGAGCGCGCGCCCGCGCGGATGCTGCTGTGCGCTGTATCTGAGATCATCGCCGTT	303	Qy	381	TrrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys	400
Qy	61	GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrrpGlnLysMetGluLys	80	Db	1264	TGGCAAGTCGCTGTGGGAAGTTTCTGGCCATCAATGCCACAAACATGTCTCTGTCTGT	132
Db	304	GAGGAAACAGACGTTTACGCGGAAACATCAAGGCAGTGGAAAATGGCAGAAAATGGA	363	Qy	401	ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu	420
Qy	81	ProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrrpLysTrrpAla	100	Db	1324	CGCGGAGAGCCCAAGGGGCTCTCCCGGCTGCCCATTTGGGAGACGGGTCTTCTGACCTC	138
Db	364	CCTTACGCTTTTACAGTTCACTGTGTAAGAGAGCAGCAGCCGCTGGAAAGTGGCGG	423	Qy	421	IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn	440
Qy	101	GlnValThrPheTrrpCysProGluGluGlnLeuCysHisLeuTrrpLeuGlnThrLeuArg	120	Db	1384	ATCCTCATCCGGAATGCTCCAAGTTCAATTTTCTGAGATTTTCTCATCAGGCACACCAAC	144
Db	424	CAGGTGACTTCTGCTGTCCAGAGGAGCAGCTGTGTCATCTTGTGGCTGCAGACCTGCGG	483	Qy	441	GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe	460
Qy	121	GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe	140	Db	1444	CAGCAGGACCACTTTCATCTTCACTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGATT	150
Db	484	GAGATGCTGGAGAAGCTGACGCTCCAGACCAAGCAATTTACTGGTATTTATCAACCCGTTT	543	Qy	461	ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe	480
Qy	141	GlyGlyLysGlyGlnGlyLysArgTrrpGluArgLysValAlaProLeuPheThrLeu	160	Db	1504	ACGTGGAAGCACATGGAGGATGAGCAGCAGCCTCAAGGAGGGGGGGAAGAGCGCTTT	156
Db	544	GGAGGAAAGAGCAAGCGGATATGAAAGAAAGTGGCACCACTGTTTCACTTCACTTA	603	Qy	481	GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrrpAsn	500
Qy	161	AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu	180	Db	1564	GGGCACATTTTGCAGCAGCCACCCCTCTGCTGCTGCACCGTCTCCAAACAGCTCTCGGAAC	162
Db	604	GCCTCCATCACCTGACATCATGCTTACTGAACATGCTAATCAGGCCAAGGAGACTCTG	663	Qy	501	CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal	520
Qy	181	TyrGluIleAsnIleAspLysTrrpAspGlyIleValCysValGlyLysGlyMetPhe	200	Db	1624	TGCGATGGGAGGCTCTGCACAGCCCTGCCATCGAGGTGAGGTCTCACTGCCAGCTGGTT	168
Db	664	TATGAGATTAACATAGACAAATACAGCGCATCGTCTGTGTCGGCGGAGATGTTATGTT	723	Qy	521	ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer	537
Qy	201	SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn	220	Db	1684	CGACTCTTTGCAGCAGGAATTTGAAGAGAAATCCGAAGCCAGACTCACACAGC	1734
Db	724	AGCGAGGTGCTGACGGTCTGATTGGGAGGACGACAGAGGCGCGGGGTGCAGCAGAAC	783	RESULT 7			
Qy	221	HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAlaGlySer	240	ADN62844			
Db	784	CACCCCGGGCTGTGCTGCTCCCGTAGCTCCCGATTGGAAATCATTTCCCGCAGGTCCA	843	ID	ADN62844 standard; DNA; 1740 BP.		
Qy	241	ThrAspCysValCysTrrpThrValGlyThrSerAspAlaGluThrSerAlaLeuHis	260	XX	AC ADN62844;		
Db	844	ACGGAATCGGTGTGTTACTCCACCGTGGGCACGACGACGAGAACCTCGCGCTGCAT	903	XX	01-JUL-2004 (first entry)		
Qy	261	IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu	280	XX	Human NOV9a DNA.		
Db	904	ATCGTGTGTGGGACTCGCTGGCCATGGATGTGCTCAGTCCACCAACAGCAGCACTC	963	KW	ds; gene; human; NOV9; metabolic disorder; diabetes; obesity;		
Qy	281	LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleLysAspSer	300	KW	infectious disease; anorexia; cancer; cancer-associated cachexia;		
Db	964	CTTCGCTACTCCGTGCTCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAGGACAGT	1023	KW	neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;		
Qy	301	GluLysLysArgTrrpLeuGlyLeuAlaArgTrrpAspPheSerGlyLeuLysThrPheLeu	320	KW	immune disorder; haematopoietic disorder; dyslipidaemia;		
Db	1024	GAGAGAAACCGTGGTGGTCTTGCAGATACGACTTTTTCAGGTTTAAAGACCTTCTC	1083	XX	metabolic syndrome X; wasting disorder.		
Qy	321	SerHisHisCysTrrpGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer	340	OS	Homo sapiens.		
				XX	US2004038223-A1.		
				XX	26-FEB-2004.		
				XX	01-OCT-2002; 2002US-00262511.		
				XX	02-OCT-2001; 2001US-0326483P.		
				PR	05-OCT-2001; 2001US-0327435P.		
				PR	05-OCT-2001; 2001US-0327449P.		
				PR	09-OCT-2001; 2001US-0327917P.		
				PR	09-OCT-2001; 2001US-0328029P.		
				PR	09-OCT-2001; 2001US-0328044P.		
				PR	09-OCT-2001; 2001US-0328056P.		

PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 18-OCT-2001; 2001US-0330309P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0339266P.  
 PR 24-OCT-2001; 2001US-033629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 17-APR-2002; 2002US-0373280P.  
 PR 19-APR-2002; 2002US-0373815P.  
 PR 19-APR-2002; 2002US-0373817P.  
 PR 19-APR-2002; 2002US-0373826P.  
 PR 19-APR-2002; 2002US-0373884P.  
 PR 22-APR-2002; 2002US-0374977P.  
 PR 16-MAY-2002; 2002US-0381037P.  
 PR 16-MAY-2002; 2002US-0381038P.  
 PR 16-MAY-2002; 2002US-0381042P.  
 PR 17-MAY-2002; 2002US-0381642P.  
 PR 28-MAY-2002; 2002US-0381656P.  
 PR 29-MAY-2002; 2002US-0383831P.  
 PR 25-JUN-2002; 2002US-0391335P.  
 XX  
 PA (SMIT/) SMITHSON G.  
 PA (MILL/) MILLET I.  
 PA (PEYM/) PEYMAN J A.  
 PA (KEKU/) KEKUDA R.  
 PA (JUJ/) JU J.  
 PA (LILL/) LI L.  
 PA (GUOX/) GUO X.  
 PA (PATT/) PATTURAJAN M.  
 PA (SPYT/) SPYTEK K A.  
 PA (EDIN/) EDINGER S R.  
 PA (ELLE/) ELLERMAN K.  
 PA (MALY/) MALYANKAR U M.  
 PA (ORTT/) ORT T.  
 PA (GORM/) GORMAN L.  
 PA (ZERH/) ZERHUSEN B D.  
 PA (ANDE/) ANDERSON D W.  
 PA (ZHON/) ZHONG M.  
 PA (CATW/) CATTERTON E.  
 PA (JIW/) JI W.  
 PA (MILL/) MILLER C E.  
 PA (RAST/) RASTELLI L.  
 PA (STON/) STONE D J.  
 PA (PENA/) PENA C E A.  
 PA (SHEN/) SHENOY S G.  
 PA (SHIM/) SHIMKETS R A.  
 PA (ROTH/) ROTHENBERG M E.  
 PA (LEAC/) LEACH M D.  
 PA (AGEE/) AGEE M L.  
 PA (BERG/) BERGHS C.  
 PA (DIPI/) DIPIPPO V A.  
 PA (EISE/) EISEN A.  
 PA (GANG/) GANGOLLI E A.  
 PA (RIEG/) RIEGER D K.  
 PA (SPAD/) SPADERNA S K.  
 XX  
 PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;  
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;  
 PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;  
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;  
 PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;  
 PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;  
 XX  
 DR WPI; 2004-213931/20.  
 DR P-PSDB; ADN62845.  
 XX  
 PT Isolated NOVX polypeptides and nucleic acids, useful for preventing,  
 PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.  
 XX  
 PS Claim 20; SEQ ID NO 39; 395pp; English.  
 XX

CC The invention relates to isolated NOVX polypeptides and polynucleotides.  
 CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or  
 CC treat a medical condition in human related to the aberrant expression and  
 CC activity of NOVX polypeptides. For example, NOVX polypeptides and  
 CC polynucleotides may be used to treat disorders associated with decreased  
 CC expression or activity of NOVX by supplementing the patient our  
 CC production or to rectify mutations. Conversely, antisense NA molecules  
 CC may be administered to down regulate expression of NOVX polypeptides by  
 CC binding with the cells own genes and preventing their expression. NOVX  
 CC polynucleotides and complementary sequences may also be used as DNA  
 CC probes in diagnostic assays to detect and quantitate the presence of  
 CC similar sequences in samples, and so which patients may be in need of  
 CC restorative therapy. NOVX polypeptides may also be used as antigens in  
 CC the production of antibodies and in assays to identify modulators  
 CC (agonists and antagonists) of the expression and activity of NOVX. The  
 CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be  
 CC used to modulate NOVX polynucleotide expression and activity of NOVX  
 CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as  
 CC diagnostic agents for detecting the presence of NOVX in samples. NOVX  
 CC polypeptides and polynucleotides may be used in this way to prevent,  
 CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious  
 CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative  
 CC disorders, Alzheimer's Disease, Parkinson's disorder, immune disorders,  
 CC haematopoietic disorders, and the various dyslipidaemias, metabolic  
 CC disturbances associated with obesity, the metabolic syndrome X and  
 CC wasting disorders associated with chronic diseases and various cancers.  
 CC They may also be used as antibacterial agents. The present sequence  
 CC represents DNA encoding a human NOVX protein.  
 XX  
 SQ Sequence 1740 BP; 388 A; 477 C; 530 G; 345 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 9,61e-251 Length: 1740  
 Score: 2843.00 Matches: 535  
 Percent Similarity: 99.63% Conservative: 1  
 Best Local Similarity: 99.44% Mismatches: 2  
 Query Match: 98.44% Indels: 2  
 DB: 12 Gaps: 0

US-10-631-958-10 (1-537) x ADN62844 (1-1740)

Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnGlnArg 20  
 Db 76 ATGGGGCGCAGCGGGCGGGCGGAGCCGCTGCAATCCGCTGCTGGTGAAGCAGCAGCGC 135  
 Qy 21 CysAlaValSerLeuGluProAlaAraGAlaLeuLeuArgTrpTrpArgSerProGlyPro 40  
 Db 136 TGGCGCCGTGAGCCCTGGAGCCCGCGGGGCTCTGCTGCGCTGGTGGCGAGCCCGGGGCC 195  
 Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal 60  
 Db 196 GGAGCGCGCGCCCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 255  
 Qy 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80  
 Db 256 GAGGAAACAGACGCTTCACGGGAAACATCAAGGCGAGTCGAAAAATGGCAGAAAAATGGAAGA 315  
 Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTyrLysTrpAla 100  
 Db 316 CCTTACCCCTTTTACAGTTCTACTGTGTAAAGAGACGACGACGCCACCCCTGGAAGTGGCG 375  
 Qy 101 GlnValThrPheTrpCysProGluGlnLeuLysHisLeuTrpLeuGlnThrLeuArg 120  
 Db 376 CAGGTGACTTTCTGGTGTGTGTCAGAGGAGCAGCTGTGTCTCACTTGTGTGCTGCAGACCTGCG 435  
 Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140  
 Db 436 GAGATGTGGAGAGCTGACGTCACAGACCAAGCAATTTACTGTGTTATTTATCAACCCGTTT 495  
 Qy 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160  
 Db 496 GGAGGAAAGACAAAGCAAGCGGATATATGAAAGAAAAGTGGCCACCACCTGTTTACCTTA 555



QY 161 AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaIleValThrLeu 180  
Db 556 GCCTCCATCACCACCTGACATATCGTTACTGAACATGCTAATTCAGGCCAAGAGACTCTCG 615  
QY 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAsp-GlyMetPh 200  
Db 616 TATGAGATTACATAGACAAATACACGGCAT-GTCTGTGTGCGGGAGATCGGTATGTT 674  
QY 200 eSerGluValLeuHisGlyLeuIleGlyValThrGlnArgSerAlaGlyValAspGlnAs 220  
Db 675 CAGCGAGGTGCTGCAGCGTCTGATTGGGAGGACGACAGAGGCGCGGGTGCACCGAA 734  
QY 220 nHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySe 240  
Db 735 CCACCCCGGGGTGCTGCTGCCAGTAGCCTCCGGATTGGAATCATTTCCGCGAGGTC 794  
QY 240 rThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHi 260  
Db 795 AACGGACTGCGTGTGTACTCCACCGTGGGCACGACGCGGAAACCTGGCGCTGCA 854  
QY 260 sIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLe 280  
Db 855 TATCGTTGTGGGACTGCTGGCCATGATGTGCTCCTCAGTCCACACCACAGCACACT 914  
QY 280 uLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSe 300  
Db 915 CTTTCGCTACTCTCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 974  
QY 300 rGluIleValArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLe 320  
Db 975 TGAGAAGAAACCGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1034  
QY 320 uSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySe 340  
Db 1035 CTCCACCACTGCTATGAAGGACAGTGTCTCTCTCCCTGCACAAACACCGTGGATC 1094  
QY 340 rProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGl 360  
Db 1095 TCCAAAGGATAGAAAGCCCTGCGGGCAGGATGCTTTGTTGACGGCAAGCAAGCAGCA 1154  
QY 360 nLeuGluGluGlnLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGl 380  
Db 1155 GCTGAGGAGGAGCAGAAAGCACTGTATGTTTGAAGCTGCGAGGACGTGGAGGA 1214  
QY 380 uTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCy 400  
Db 1215 GTGGCAAGTCTGCTGGGAAGTTTCTGGCCATCAATGCCAACATGCTCTGTGCTTG 1274  
QY 400 sArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLe 420  
Db 1275 TGGCCGAGCCCCCAGGGGCTCTCCCGGCTGCCCCACTTGGGAGACGGGTCTTCTGACCT 1334  
QY 420 uIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAs 440  
Db 1335 CATCTCTCATCGGAAATGCTCAGGTTCATATTTCTGAGATTTCATCAGGCACACCAA 1394  
QY 440 nGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPh 460  
Db 1395 CCAGCAGGACCAATTTGATCTTCTTGTGAGTTTATCGGTCAAGAAATTCAGATT 1454  
QY 460 eThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPh 480  
Db 1455 TACGTTCGAAGCAGCATGGAGATGAGCAGCGACCTCAAGGAGGGGGGAGAGCGCTT 1514  
QY 480 eGlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAs 500  
Db 1515 TGGGCACATTTGAGCAGCAGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1574  
QY 500 nCysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVa 520  
Db 1575 CTGCCAGGGGAGGTCTGTCACAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1634  
QY 520 lArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 537

Db 1635 TCGACTCTTTGCACGAGGATTTGAAGAGATCCGAAGCCAGACTCACACAGC 1686  
RESULT 8  
AAD14426  
ID AAD14426 standard; cDNA; 1840 BP.  
XX  
AC AAD14426;  
XX  
DT 01-NOV-2001 (first entry)  
XX  
DE Human sphingosine kinase (SphK) cDNA #2.  
XX  
KW Human; sphingosine kinase; SphK; restenosis; ischaemia; gene therapy;  
antisense therapy; cancer; sphingolipid; signalling molecule; apoptosis;  
cytostatic; tumour necrosis factor-alpha; TNF; atherosclerosis; lymphoma;  
leukaemia; vasotropic; cell proliferative disorder; vascular disease; ss.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 109..1524  
FT /\*tag= a  
FT /product= "Human sphingosine kinase (SphK) protein #2"  
XX  
PN WO200160990-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 14-FEB-2001; 2001WO-US004789.  
XX  
PR 14-FEB-2000; 2000US-0182360P.  
PR 22-MAR-2000; 2000US-0191261P.  
XX  
PA (CURA-) CURAGEN CORP.  
PA (GETH ) GENENTECH INC.  
XX  
PI Raetelli L;  
XX  
DR WPI; 2001-514770/56.  
XX  
P-PSDB; AAE07884.  
XX  
PT An isolated sphingosine kinase polypeptide useful for treating a SphK-  
associated disorder especially cancer, restenosis or ischemia in a human.  
XX  
PS Claim 8; Page 94-95; 107pp; English.  
XX  
CC The present invention relates to sphingosine kinase (SphK) polypeptides  
and nucleic acids encoding them. SphK is useful for treating a SphK-  
associated disorder especially cancers such as leukaemia, lymphoma,  
ovarian, breast, lung, colon, testicular, stomach and skin,  
atherosclerosis, restenosis or ischaemia and cell proliferative disease  
or disorder associated with vascular diseases. SphK gene is used in gene  
therapy and antisense-therapy. Sphingolipids serving as signalling  
molecules, have recently emerged as regulators of cell growth,  
differentiation, diverse cell phenotypes and cell death. Activation of  
SphK by tumour necrosis factor (TNF)-alpha inhibits apoptosis in human  
endothelial cells. The present sequence is human sphingosine kinase  
(SphK) cDNA  
XX  
SQ Sequence 1840 BP; 459 A; 468 C; 503 G; 410 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 6,01e-233 Length: 1840  
Score: 2649.50 Matches: 495  
Percent Similarity: 97.83% Conservative: 1  
Best Local Similarity: 97.63% Mismatches: 0  
Query Match: 91.74% Indels: 11  
DB: 5 Gaps: 1  
US-10-631-958-10 (1-537) x AAD14426 (1-1840)

QY 42 AlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleIleAlaValGlu 61



[illegible][illegible]

CC sphingosine to form sphingosine 1-phosphate. The polynucleotide was  
CC isolated from an Hela cDNA library by PCR amplification. The invention  
CC provides polynucleotides (see AAA50508-10) and polypeptides (see AAY96057  
CC -59) for the human sphingosine kinase (SK) homologues SKA, SKB and SKC.  
CC The polynucleotides may be used as hybridization probes, in the  
CC construction of PCR primers for chromosome and gene mapping, in the  
CC recombinant production of SKA, SKB and SKC, and in the generation of  
CC antisense DNA or RNA. They can be used to detect inflammation or disease  
CC associated with abnormal levels of SK expression, or to detect  
CC differences in gene sequence between normal and carrier or affected  
CC individuals. Host cells expressing SK can be used in drug screening.  
CC Human SK specific antibodies, inhibitors, ligands or their analogues are  
CC useful as bioactive agents to treat inflammation or disease including  
CC viral, bacterial or fungal infections, allergic responses, mechanical  
CC injury associated with trauma, hereditary diseases, lymphoma or  
CC carcinoma, and other conditions with activate the genes of kidney, lung,  
CC heart, lymphoid or tissues of the nervous system  
XX

SQ Sequence 4231 BP; 1022 A; 1021 C; 1086 G; 1067 T; 0 U; 35 Other;

Alignment Scores:  
Pred. No.: 1.05e-226 Length: 4231  
Score: 2587.00 Matches: 481  
Percent Similarity: 99.79% Conservative: 1  
Best Local Similarity: 99.59% Mismatches: 1  
Query Match: 89.58% Indels: 0  
DB: 3 Gaps: 0

US-10-631-958-10 (1-537) x AAA50510 (1-4231)

QY 55 SerGluIleAlaValGluThrAspValHisGlyLysHisGlnGlySerGlyLys 74  
DB 2 TCTGAGATCATCGCCGTTTGAGGAAACACACGCTTACCGGAAACATCAAGGCGAGTGAATA 61

QY 75 TrpGlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgArg 94  
DB 62 TGGCAGAAATGGAAGAAGCCCTACGCTTTTACAGTTCTACTGTGTAAAGAGACGACGCG 121

QY 95 HisArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLysCysHisLeu 114  
DB 122 CACCGCTGGAAGTGGCGCAGGTACTTCTGCTGTCCAGAGGAGCAGCTGTGTCACTTG 181

QY 115 TrpLeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeu 134  
DB 182 TGGCTGCAGACCCCTCGGAGATGCTGGAGAAGCTGACGTCAGACCAACATTTACTG 241

QY 135 ValPheIleAsnProPheGlyLysGlyGlnGlyLysArgGlyLeuThrGluArgLysVal 154  
DB 242 GTATTATACACCGTTTGAGGAAAGGACACAGCAGCGGATATATGAAGAAAGAGTG 301

QY 155 AlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsn 174  
DB 302 GCACCACTGTTTCACTTACCTCCATCACCAGCTGACATCATCTGTACTGAACATGCTAAT 361

QY 175 GlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysVal 194  
DB 362 CAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGACGGCATCTGTGTGC 421

QY 195 GlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyValThrGlnArgSer 214  
DB 422 GCGGAGATGTATGTTTCAGCGAGGTGTGCACGGTCTGATTTGGAGACACAGAGGAGC 481

QY 215 AlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGly 234  
DB 482 GCGGGGTCCAGCAGAACACCCCGGCTGTGCTGGTCCCAGTAGCTCCGGATTGGA 541

QY 235 IleIleProAlaGlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAla 254  
DB 542 ATCATTCGCGAGGGTCAACGAGCTGCGTGTGTACTCCACCGTGGGCACGAGCGCA 601

QY 255 GluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspValSerVal 274  
DB 602 GAAACCTCGGCGCTGCATATCTGTTGGGGACTCGCTGGCCATGGATGTCTCTCAGTC 661

QY 275 HisHisAsnSerThrLeuLeuAArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGly 294  
DB 662 CACCACAACAGCACACTCTTTCGCTACTCCGTGTCCCTGCTGGGCTACGGCTTCTACGGG 721

QY 295 AspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaAArgTyrAspPheSer 314  
DB 722 GACATCATCAAGGACAGTGAGAAACGGTGGTGTTCGTCAGATACGACTTTTCA 781

QY 315 GlyLeuLysThrPheLeuSerHisCysTyrGluGlyThrValSerPheLeuProAla 334  
DB 782 GGTTTAAAGACCTTCTCTCCACCACTGCTATGAAGGACAGTGTCTCTCTCCCTGCA 841

QY 335 GlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCys 354  
DB 842 CAACACACGGTGGGATCTCCAAGGAGATAGGAAGCCCTGCCCGGAGGATGCTTGTGTC 901

QY 355 ArgGlnSerLysGlnGlnLeuGluGluGlnLysValAlaLeuTyrGlyLeuGluAla 374  
DB 902 AGGCAAGCAAGCAGCAGCTGGAGGAGCAGAAAGAACGACTGTATGGTGTGGAGCT 961

QY 375 AlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThr 394  
DB 962 GCGGAGGACGTTGGAGGAGTGGCAAGTCTCTGTGGGAAGTTTCTGGCCATCAATGCCACA 1021

QY 395 AsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGly 414  
DB 1022 AACATGTCTGTGCTTGTGCGCGGAGGCCACAGGGGCTCTCCCGGCTGCCACTTGGGA 1081

QY 415 AspGlySerSerAspLeuIleLeuLeuArgLysCysSerArgPheAsnPheLeuArgPhe 434  
DB 1082 GACGGGTCTCTGACCTCATCTCTCATCCGGAATATGCTCCAGGTTCATATTTCTGAGATT 1141

QY 435 LeuIleArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArg 454  
DB 1142 CTCTATCAGGCACACCAACAGCAGGACGACGTTTCACTTCTCTTTGTTGAAGTTATGCG 1201

QY 455 ValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGlu 474  
DB 1202 GTCAAGAAATTCAGTTTACGTCGAAGCACATGAGGATGAGGACAGCAGCTCAAGGAG 1261

QY 475 GlyLysLysLysArgPheGlyHisIleCysSerSerHisProSerCysCysThrVal 494  
DB 1262 GGGGGGGAAGAGCGCTTTGGGCACATTTGACAGCAGCAGCCCTCTCTGTCGACCGTC 1321

QY 495 SerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArg 514  
DB 1322 TCCAAACAGCTCTGGAACCTGGCAGCGGAGGCTCTGCACAGCCCTGCCATCGAGGTGAGA 1381

QY 515 ValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnProLysProAsp 534  
DB 1382 GTCCACTGCCAGCTGGTTCGACTCTTTGCACGAGATTTGGAAGAGAAATCCGAAGCCAGAC 1441

QY 535 SerHisSer 537  
DB 1442 TCACACAGC 1450

RESULT 10  
ADS10370  
ID ADS10370 standard; DNA; 4702 BP.  
XX  
AC ADS10370;  
XX  
DT 16-DEC-2004 (first entry)  
XX  
DE Human therapeutic DNA - SEQ ID 607.  
XX  
KW antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnerary;  
KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;  
KW aplastic anaemia; cancer; wound healing; gene therapy; ds; gene.  
XX  
OS Homo sapiens.  
XX

PN WO2004080148-A2.  
XX 23-SEP-2004.  
XX 30-SEP-2003; 2003WO-US030720.  
XX 02-OCT-2002; 2002US-0416186P.  
XX (NUVE-) NUVELO INC.  
XX Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y;  
XX Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AU, Weng G, Zhou P;  
XX WPI; 2004-668857/65.  
XX P-PSDB; ADS11054.  
XX New polynucleotide, useful in preparing a composition for diagnosing or  
XX treating inflammatory, neurodegenerative or stem cell disorders, e.g.,  
XX aplastic anemia or cancer for promoting wound healing.  
XX Claim 1; SEQ ID NO 607; 718pp; English.  
XX The invention relates to a novel isolated polynucleotide and the encoded  
XX polypeptide. The molecules of the invention demonstrate antiinflammatory,  
XX neuroprotective, antianaemic, cytostatic and vulnerary activities and may  
XX be useful in preparing a composition for diagnosing or treating  
XX inflammatory, haematopoietic, immune, neurodegenerative or stem cell  
XX disorders, such as aplastic anaemia or cancer, as well as for promoting  
XX wound healing. The molecules may also be utilised during gene therapy  
XX procedures. The current sequence is that of a human therapeutic DNA of  
XX the invention. The current sequence is not shown explicitly within the  
XX specification but can be accessed from the WIPO web-site.  
XX Sequence 4702 BP; 970 A; 1226 C; 1364 G; 1118 T; 0 U; 24 Other;  
SQ  
Alignment Scores:  
Pred. No.: 7.86e-218 Length: 4702  
Score: 2491.00 Matches: 493  
Percent Similarity: 65.87% Conservative: 1  
Best Local Similarity: 65.73% Mismatches: 1  
Query Match: 86.25% Indels: 256  
DB: 13 Gaps: 3  
US-10-631-958-10 (1-537) x ADS10370 (1-4702)  
Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnArg 20  
Db 818 ATGGGGGGCGACGGGGGGCGGGAGCCGCTGCAATCCGCTGTGGGTGAAGCAGCAGCGC 877  
Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40  
Db 878 TGGCCGCTGAGCTGGAGCCCGCGGGGCTCTGCTGGCTGTGGCGGAGCCCGGGGCC 937  
Qy 41 GlyAlaGlyAlaProGlyAla----- 47  
Db 938 GGAGCGGGCGCCCGGGGCG-GGCCTCACCAGGCCAGCCCTCTTCCAGACTCACCT 996  
Qy 47 ----- 47  
Db 997 AGCACTGGTTATTTTGGCAGTAGTTTCCAAGACTGGGCGAGCGTTTACACTTCTGTTAGC 1056  
Qy 47 ----- 47  
Db 1057 TGGTTTCTCTGGCACTGCTCCAGCAGGGAACGGCGCTGGTACTGCCAGGGGCGCTGCACCC 1116  
Qy 47 ----- 47  
Db 1117 ACCCTGTGCCCCCCTAGCTCAGAGCTCCCGAGGCCAGAGCTGGAGATAAACCCAGCTCC 1176  
Qy 47 ----- 47  
Db 1177 CAGCTGTGACTCTGCGACGCTGGGCTCTCCAGGCCAGCGGGGCGCTGCTGTGGATCTCTG 1236

Qy 47 ----- 47  
Db 1237 GTGTCCCGGTTCACTGCGCTGTTTCTTGTATCTTGTCCAGAGCTATCAGTCATAGCGGA 1296  
Qy 47 ----- 47  
Db 1297 CGGAGGCTGATGACGATTTCCTCCGCGACCTGA CAGCATGAGTGGGACCGCGGGTGTTC 1356  
Qy 48 -----AspAlaCysSerValProVal 54  
Db 1357 TGTGTTGAGATGCCAGCTGAAGGACGTGGTGTGTTTACAGATGCTGCTCTGTGCTGTA 1416  
Qy 55 SerGluIleIleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLys 74  
Db 1417 TCTGAGATCATCGCGTTGAGGAAACAGACGTTTACGGGAAACATCAAGCAGCTGGAAAA 1476  
Qy 75 TrpGlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArgArg 94  
Db 1477 TGGCAGAAAATGGAAAAGCCCTTACGCTTTTACAGTTTCACTGTCTAAAGAGAGACGACGG 1536  
Qy 95 HisArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCysHisLeu 114  
Db 1537 CACCGCTGGAAAGTGGCGCAGGTGACTTCTGCTGTCAGAGGAGCAGCTGTGTCACTG 1596  
Qy 115 TrpLeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeu 134  
Db 1597 TGGCTGCAGACCCCTGCGGGAGATGCTGGAGAAGCTG----- 1632  
Qy 135 ValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysVal 154  
Db 1632 ----- 1632  
Qy 155 AlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsn 174  
Db 1633 -----ATTACTGAACATGCTAAT 1650  
Qy 175 GlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysVal 194  
Db 1651 CAGGCCAAGGAGACTCTGTATGAGATTAACATAGACAAATACGACGCGCATCGTGTGTGC 1710  
Qy 195 GlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSer 214  
Db 1711 GCGCGAGATGTTATGTTTCAGCGAGGTGCTGCACGGTCTGATGGGAGACGCGAGAGGC 1770  
Qy 215 AlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGly 234  
Db 1771 GCGGGGTCCACAGAACACACCCCGGGGCTGTGCTGCTCCAGTAGCTCCGATTTGA 1830  
Qy 235 IleIleProAla----- 238  
Db 1831 ATCATTTCCGCGAGCTTTTGCAATGACCTGGCAGGGGACGAGGTGTCTGCTCTCTGGC 1890  
Qy 238 ----- 238  
Db 1891 CCTGTGTGTCGCCCGAGGGTGGCGCATGGTGACACTTTTCACTCTCGTCACAGCTCTG 1950  
Qy 238 ----- 238  
Db 1951 GGATGTGAGCACCGCAGCTCATCCCCATTTTATGGATGAACAGCAGGAGACTGGGGGAGCAT 2010  
Qy 239 -----GlySerThrAspCysValCysTyrSer 247  
Db 2011 GTGGGCCCCGGTGAGAACGCTGTGGCTTGGACGGGTCAACGGGACTCGTGTGTACTCC 2070  
Qy 248 ThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeu 267  
Db 2071 ACCGTGGGCGCCAGCAGCAGAAACCTCGCGCTGCATATCGTTGTGTGGAGCTCCGCTG 2130  
Qy 268 AlaMetAspValSerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeu 287  
Db 2131 GCCATGATGTGTCTCAGTCCACCACACAGCACACTCTCTCGCTACTCCGTGCTCCCTG 2190  
Qy 288 LeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGly 307

```
Db 2191 CTGGGCTACGGCTTCTACGGGACATCATCAGGACAGTGAGAGAAACGGTGTGGGT 2250
Qy LeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCyeTyrGluGly 327
Db 2251 CTGGCAGATACGACTTTTCAGGTTTAAAGACCTTCTCTCCACCACTGCTATGAGGG 2310
Qy ThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysProCys 347
Db 2311 ACAGGTCTCTTCTCCCTCCGCAACACACGCGTGGGATCTCAAGGGATAGGAAGCCCTGC 2370
Qy ArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLeuGluGluGluGlnLysLys 367
Db 2371 CGGGCAGATGCTTTGTTTGCAGGCAAGCAGCAGCTGAGGAGCAGCAGAGAA 2430
Qy AlaLeuTyrGlyLeuGluAlaGluAspValGluGluTrpGlnValValCysGlyLys 387
Db 2431 GCACGTGTATGTTTGAAGCTCGCGAGCAGCTGGAGGAGTGCGAAGTCTGTGTGGGAG 2490
Qy PheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeu 407
Db 2491 TTTCTGGCCATCAATGCCACAAACATGCTGTGTGCTGTCGCGAGCCCGAGGGCCCTC 2550
Qy SerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIleLeuArgLysCysSer 427
Db 2551 TCCCGGCTGCCACTTGGGAGCGGTCTTCTGACCTCATCTCCGAAATGCTCC 2610
Qy ArgPheAsnPheLeuArgPheLeuIleArgHisThrAsnGlnGlnAspGlnPheAspPhe 447
Db 2611 AGGTTCAATTTCTGAGATTTCTCATCAGGCACACCAACAGCAGCAGCAGTTGACTTC 2670
Qy ThrPheValGluValTyrArgValLysLysPheGlnPheThrSerLysHisMetGluAsp 467
Db 2671 ACTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGATTTCAGTCGAAGCAGCATGGAGAT 2730
Qy GluAspSerAspLeuLysGluGlyLysLysArgPheGlyHisIleCysSerSerHis 487
Db 2731 GAGGACGACCTCAGAGGGGGGGAAGCGCTTTGGGCATTTGCGACAGCCAC 2790
Qy ProSerCysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHis 507
Db 2791 CCTCTCTGCTGTCACCGTCTCCAAACAGCTCCTGGAACCTGGACGGGGAGTCTGCAC 2850
Qy SerProAlaIleGluValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIle 527
Db 2851 AGCCCTGCATCAGAGTCAGAGTCACATGCCAGCTGTTGCACTTTTGCACGAGGAAT 2910
Qy GluGluAsnProLysProAspSerHisSer 537
Db 2911 GAAGAGAAATCGAAGCCAGACTCACACAGC 2940

RESULT 11
AAS77728
ID AAS77728 standard; cDNA; 2241 BP.
AC AAS77728;
XX 13-FEB-2002 (first entry)
DT
XX DNA encoding novel human diagnostic protein #13532.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX
```

```
PR 23-AUG-2000; 2000US-00649167.
PA (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG13541.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 1; SEQ ID NO 13532; 103pp; English.
PS The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridization probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2241 BP; 453 A; 627 C; 700 G; 461 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 2,73e-205 Length: 2241
Score: 2349.50 Matches: 471
Percent Similarity: 64.50% Conservative: 5
Best Local Similarity: 63.82% Mismatches: 8
Query Match: 81.35% Indels: 255
DB: Gaps: 5

US-10-631-958-10 (1-537) x AAS77728 (1-2241)
Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnArg 20
Db 83 ATGGGGCGAGCGGGGGCGGAGCCGCTGCAATCCGTGCTGGTGAAGCAGCAGCGC 142
Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40
Db 143 TGCGCCGTGAGCCTGGAGCCCGCGCGGCTCTGCTGCGCTGGTGGCGAGCCCGGGGCC 202
Qy 41 GlyAlaGlyAlaProGlyAla----- 47
Db 203 GGAGCCGGCGCCCGCGCGCGGCGGCGGCGGCTCACCAGGCCAGCCCTCTTCCCAGACTCACC 261
Qy 47 ----- 47
Db 262 AGCACTGTTATTTTGCAGTATGTTTCCAAGACTGGCGAGCGGTTTACACCTTCTGTTTAGC 321
Qy 47 ----- 47
Db 322 TGGTTTCTCTGGCACTGCTCCAGCAGGGAACGGCGCTGGTACTGCCAGGGGCCCTGCACCC 381
Qy 47 ----- 47
Db 382 ACCCTGCTGCCCATAGCTCAGACGTCCTCCGAGGCCAGAGCTGGAGATAAACCCAGCTCC 441
```

```
Qy 47 ----- 47
Db 442 CAGCTGCTGACTCTGCAGCTGGCCTCTCCAGCCAGCGGGCCTGCTGCTGGGATCTCTG 501
Qy 47 ----- 47
Db 502 GTGTCCCGGTTTCACTGCTGTTTCTTGTAATCTTTGTCAGAGCTATCAGTCATAGCGGGA 561
Qy 47 ----- 47
Db 562 CGGAGGCTGATCAGGTTTCCCGCGACCTGACAGCATGGAGTGGGACCGCGGTGTTTC 621
Qy 48 ----- AspAlaCysSerValProVal 54
Db 622 TGTGTTGAGATGCCAGCTGAAGGACGTGGTGTGTTTTACAGATGCTGCTCTGTGCTGTA 681
Qy 55 SerGluIleIleAlaValGluGluThrAspValHisGlyLysHisGlnGlySerGlyLys 74
Db 682 TCTGAGATCACTGCCGCTTAGGAAACACAGCGTTTCAACGGAACATCAAGGCAAGTGGAAA 741
Qy 75 TrpGlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArgAlaArg 94
Db 742 TGGCAGAAATGGAAAGCCCTTACGCTTTTACAGTTTCACTGTGTAAAGAGACGACGCG 801
Qy 95 HisArgTrpLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuLysHisLeu 114
Db 802 CACCGCTGGAAGTGGCGCAGGTGACTTCTGTGTGTCAGAGGAGCAGCTGTGTCACTTG 861
Qy 115 TrpLeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeu 134
Db 862 TGGCTGCAGACCTTGGCGGAGATGCTGAGAAGCTG----- 897
Qy 135 ValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysVal 154
Db 897 ----- 897
Qy 155 AlaProLeuPheThrLeuAlaSerIleThrAspIleIleValThrGluHisAlaAsn 174
Db 898 -----ATTACTGAACATGCTAAT 915
Qy 175 GlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysVal 194
Db 916 CAGGCCAAGGAGACTCTGTATGAGATTAAACATAGACAAATACGACGCGCATCGTGTGTC 975
Qy 195 GlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSer 214
Db 976 GCGCGAGATGGTATGTTTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGACGAGGAGC 1035
Qy 215 AlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGly 234
Db 1036 GCGCGGTTCGACCAAGAACCCACCGCGGCTGTGTCGTCCTCCAGTAGCCTCCGATTTGGA 1095
Qy 235 IleIleProAla-----GlySerThr 241
Db 1096 ATCATTCGCGAGAGCATGTGGGCCCGGTGAGAACCGTGTGGCTTGAGACGGGTCAACG 1155
Qy 242 AspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIle 261
Db 1156 GACTGCGGTGTGTACTCCACCGTGGGACCACGACGACGAGAAACCTCGGCGCTGCATATC 1215
Qy 262 ValVal----- 263
Db 1216 GTTGTGTGCTGTGCCCCGAGCTCGGAAACACCACCGCATCCGCCATCTGCCCTGTGGC 1275
Qy 263 ----- 263
Db 1276 AGTGGCACGGGCAGCTGTGCTGGGTGTCAGCTGTGGAAACGCATGCTCTGTGAGGCC 1335
Qy 264 -----GlyAspSerLeuAlaMetAspValSer 273
Db 1336 TCGAGGCTTCAGTCCAGGATGCAGAGCCCCGGGACTCGCTGGCCATGGATGTGTCTTCA 1395
```

```
Qy 274 ValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyr 293
Db 1396 GTCCACCAACAGCACACTCTTGGCTACTCCGTTGTCCTGTGGGTACTACGGCTTCTAC 1455
Qy 294 GlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPhe 313
Db 1456 GGGGACATCATCAAGGACAGTGAAGAAACGGTGGTGGTCTTGGCAGATACGACTTT 1515
Qy 314 SerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuPro 333
Db 1516 TCAGGTATAAAGACTTCTCTCCACCATGCTATGAAGGACAGTGTCTCTTCTCCCT 1575
Qy 334 AlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheVal 353
Db 1576 GCACACACACGCTGGGATCTCCAAGGAATAGGAAGCCCTGCCGGGAGGATGCTTTGTT 1635
Qy 354 CysArgGlnSerLysGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGlu 373
Db 1636 TGCAGGCAAAAGCAAGCAGCAGCTGGAGGAGGAGCAGAAAGAACACTGTATGTTTGGAA 1695
Qy 374 AlaAlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAla 393
Db 1696 GCTGCGGAGGAGCTGGAGGAGTGGCAAGTCTCTGTGGGAAGTTTCTGGCCATCAATGCC 1755
Qy 394 ThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeu 413
Db 1756 ACAACACATGTCCTGTGCTTGC CGGAGGCCCCAGGGGCTCTCCCGGCTGCCACTTG 1815
Qy 414 GlyAspGlySerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArg 433
Db 1816 GSAGACGGGTCTTCTGACCTCATCTCCGAAATGCTCCAGGTTCAATTTTGAAGTTAT 1875
Qy 434 PheLeuIleArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyr 453
Db 1876 TTTCTCATGAGGCACACCAACAGCAGACAGTTTGACTTCACTTTTGTGAAGTTAT 1935
Qy 454 ArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLys 473
Db 1936 CGCGTCAAGAAATTCCTGTTTACGTCGAGCACATGAGGAGTGAAGACGACCTCAAG 1995
Qy 474 GluGlyLysLysArgPheGlyHisIleCysSerSerHisProSerCysCysThr 493
Db 1996 GAGGGGGGGAAGAACGCTTTGGGCACATTTGACGACGACCCCTCTGCTGCTGCACC 2055
Qy 494 ValSerAsnSerSerTrpAsnCysAspGlyValLeuHisSerProAlaIleGluVal 513
Db 2056 GTCTCCAAACAGCTCTGGAACTGGACGGGAGGTCTCTGCACAGCCCTGCCATCGAGTC 2115
Qy 514 ArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnPro 531
Db 2116 AGG-----GCGGTGGTGTGTCGCAAGAGGCTGTGCAAAAGGTCCCA 2154
RESULT 12
ACC70838
ID ACC70838 standard; DNA; 3975 BP.
XX
AC ACC70838;
XX
XX
DT 20-NOV-2003 (first entry)
XX
DE Human Sphingosine kinase 4 coding sequence.
XX
KW Human; enzyme; haemostatic; sphingosine kinase 4; SPHK4;
XX platelet transfusion; platelet stabiliser; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 1..1251
FT /*tag= a
FT /product= "Sphingosine kinase 4"
XX
FN W02003031627-A1.
```

XX 17-APR-2003.  
 XX 28-SEP-2001; 2001WO-JP008537.  
 XX 28-SEP-2001; 2001WO-JP008537.  
 XX (HOKK-) HOKKAIDO TECHNOLOGY LICENSING OFFICE CO.  
 XX (CHBI-) CHEM BIOLOGY INST.  
 XX Igarashi Y, Kihara A;  
 XX WPI: 2003-354917/33.  
 XX P-PSDB; ABR56301.  
 XX Platelet derived polypeptides with sphingosine kinase activity for  
 XX treatment of sphingosine related disorders.  
 XX Claim 4; Page 30-31; 39pp; Japanese.  
 XX The present sequence is the coding sequence for human sphingosine kinase  
 CC 4 (SPHK4). The kinase can be used for the diagnosis and treatment of  
 CC sphingosine related disorders. The kinase can also be potentially used  
 CC for controlling toxicity of platelet transfusion and as a platelet  
 CC stabiliser  
 XX  
 SQ Sequence 3975 BP; 943 A; 978 C; 1042 G; 1012 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 3 87e-192 Length: 3975  
 Score: 2210.00 Matches: 416  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 76.52% Indels: 0  
 DB: 8 Gaps: 0  
 US-10-631-958-10 (1-537) x ACC70838 (1-3975)  
 QY 122 MetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPheGly 141  
 DB 1 ATGCTGGAGAGCTGACGTCCAGACCAAGCATTTTACTGGTATTATCAACCCGTTTGA 60  
 QY 142 GlyLysGlyGlnGlyValArgIleThrGluArgLysValAlaProLeuPheThrLeuAla 161  
 DB 61 GGAAGAAGACACAGGACGGATATATGAAGAAAGTGGACCACTGTTACCTTAGCC 120  
 QY 162 SerIleThrThrAspIleThrValThrGluHisAlaAsnGlnAlaLysGluThrLeuTyr 181  
 DB 121 TCCATCACCTGACATCATCTGTTACTGAACATGCTAATCAGGCCCAAGGACACTCTGTAT 180  
 QY 182 GluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPheSer 201  
 DB 181 GAGATTAACATAGACAAATACAGCGCATCGTCTGTGCGGGAGATGGTATGTTACG 240  
 QY 202 GluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsnHis 221  
 DB 241 GAGGTGCTGTCACGTCTGATTGGGAGGACGACAGAGAGCGCGGGGTCCAGACCAAC 300  
 QY 222 ProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySerThr 241  
 DB 301 CCCCAGGCTGCTGCTGCTCCAGTAGTACCTCCGATTTGGAATCATTTCCCGAGGGTCAACG 360  
 QY 242 AspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIle 261  
 DB 361 GACTGGGTGTGTACTCCACCGTGGGACACGACGACGAGAACTCCGGCGCTGCATATC 420  
 QY 262 ValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeuLeu 281  
 DB 421 GTTGTGGGACTCGCTGGCCATGGATGTCTCCTCAGTCCACCAACACAGACACTCTCT 480  
 QY 282 ArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleAspSerGlu 301  
 DB 481 CGCTACTCGGTGCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAGGACAGTAG 540

QY 302 LysLysArgTrpTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSer 321  
 DB 541 AAGAAACGGTGGTGGTCTTGCAGATACGACTTTTCAGGTTTAAAGACCTTCTCTCC 600  
 QY 322 HisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerPro 341  
 DB 601 CACCACCTGCTATGAAGGACAGTGTCTTCTCTCCGACCAACACAGCGTGGGATCTCCA 660  
 QY 342 ArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLeu 361  
 DB 661 AGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTGCAGGCAAGCAACAGCAGCTG 720  
 QY 362 GluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaGluAspValGluGluTrp 381  
 DB 721 GAGGAGGACGAGAAAGACACTGTATGTTTGAAGCTGCGAGGAGCTGCGAGGAGTGG 780  
 QY 382 GlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArg 401  
 DB 781 CAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCAACAAACATGTCTGTGCTTGTGCG 840  
 QY 402 ArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIle 421  
 DB 841 CGGAGCCCGGAGGCGCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTCTGACCTCATC 900  
 QY 422 LeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsnGln 441  
 DB 901 CTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCACCAACCCAG 960  
 QY 442 GlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPheThr 461  
 DB 961 CAGGACCAAGTTGACTTTCACCTTTTGTGAAGTTTATCGCGTCAAGAAATTCAGTTTACG 1020  
 QY 462 SerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPheGly 481  
 DB 1021 TCGAAGCACATGGAGGATGAGACAGCGACTCAAGGAGGGGGGAGAGACGCTTTGGG 1080  
 QY 482 HisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsnCys 501  
 DB 1081 CACATTTGCAGACGCCACCCCTCTGCTGTGTCACCCGTCTCCAACAGCTCTCGAACTGC 1140  
 QY 502 AspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuValArg 521  
 DB 1141 GACGGGAGGTCTTCGACAGCCCTGCCATCGAGTCCAGTCCACTGCCAGCTGGTTGCA 1200  
 QY 522 LeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 537  
 DB 1201 CTCCTTTCACGAGGAATTGAAGAGATCCGAGGCCAGACTCACACAGC 1248  
 RESULT 13  
 AAS77730  
 ID AAS77730 standard; cDNA; 2186 BP.  
 XX  
 AC AAS77730;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE DNA encoding novel human diagnostic protein #13534.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 XX food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US008631.  
 XX  
 PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX







Db	1443	TATCGCGTCAAGAAATTCAGTTTACGTGCGAAGCAATGAGGATGAGGACAGCGACCTC	1502
Qy	473	LysGluGlyGlyLysArgPheGlyHisIleCysSerSerHisProSerCysCys	492
Db	1503	AAGGAGGGGGGAGAGCGCTTTGGGCACATTTGCAGCAGCCACCTCTCTGCTGCTGC	1562
Qy	493	ThrValSerAsnSerSerTyrAsnCysAspGlyGluValLeuHisSerProAlaIleGlu	512
Db	1563	ACCGTCTCCACAGCTCTCTGGAACGCGACGGGAGGTCTCTGCAGCCCTGCCATCGAG	1622
Qy	513	ValArg 514	
Db	1623	GTGAGA 1628	
RESULT 14			
ID	ABL40822		
XX	AC	ABL40822;	
XX	DT	03-JUL-2002 (first entry)	
XX	DE	Human sphingosine kinase-like protein encoding cDNA.	
KW		Human sphingosine kinase-like protein; intracellular signalling; gene;	
KW		cell proliferation; apoptosis; cancer; allergy; cytotostatic; asthma;	
XX		autoimmune disease; rheumatoid arthritis; Parkinson's disease; ss.	
OS		Homo sapiens.	
XX	Key	Location/Qualifiers	
FT	CDS	2..789	
FT		/*tag= a	
FT		/product= "sphingosine kinase-like protein"	
FT		/note= "start and stop codons are not indicated"	
XX			
PN	WO200228906-A2.		
XX			
PD	11-APR-2002.		
XX			
XX	05-OCT-2001; 2001WO-EP011516.		
PR	06-OCT-2000; 2000US-0238005P.		
PR	23-AUG-2001; 2001US-0314113P.		
XX			
PA	(FARB ) BAYER AG.		
XX			
PI	Kossida S, Encinas J;		
XX			
DR	WPI; 2002-340094/37.		
DR	P-PSDB; ABB07854.		
XX			
PT	New reagent for modulating the activity of sphingosine kinase-like protein polypeptide or polynucleotide and treating cancer, asthma, allergy, an autoimmune disease, or a central or peripheral nervous system disorder.		
XX			
PS	Claim 1; Fig 1; 120pp; English.		
XX			
CC	The invention relates to a human sphingosine kinase-like protein. The polypeptide can be expressed by standard recombinant methodology. The sphingosine kinase-like protein and gene can be used to regulate intracellular signalling and consequently cell proliferation and apoptosis. Such regulation is useful for treating cancer, allergies (e.g. asthma), autoimmune diseases (e.g. rheumatoid arthritis) and central and peripheral nervous system disorders (e.g. Parkinson's disease). The present sequence represents the human sphingosine kinase-like protein encoding cDNA		
XX			
SQ	Sequence 979 BP; 237 A; 259 C; 271 G; 212 T; 0 U; 0 Other;		

Alignment Scores:  
Pred. No.: 1.04e-140 Length: 979

Score:	1640.50	Matches:	314
Percent Similarity:	96.63%	Conservative:	0
Best Local Similarity:	96.32%	Mismatches:	1
Query Match:	56.80%	Indels:	11
DB:	6	Gaps:	1
US-10-631-958-10 (1-537) x ABL40822 (1-979)			
Qy	130	ProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgIle	149
Db	2	CCAAAGCATTTTACTGTGTTATTTATCAACCCGTTTGGAGGAAAAGGACAAGCAAGCGGATA	61
Qy	150	TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIle---	168
Db	62	TATGAAAGAAAAGTGGCACCACCTGTTCCACCTTACGCTCCATCACCACCTGACATCATCGGT	121
Qy	169	-----ValThrGluHisAlaAsnGlnAlaLysGlu	178
Db	122	AACAAATTCTATGTTAACTATGTAGAAAGTAATTACTGAAACATGCTAATACGCCCAAGGAG	181
Qy	179	ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly	198
Db	182	ACTCTGTATGAGATTAAACATAGACAAATACAGCGCATCGTCTGTGTCCGCGGAGATGGT	241
Qy	199	MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp	218
Db	242	ATGTTTCAGCGAGGTGCTGCACGGTCTGATTGGGAGGACGAGGAGCGCGGGGTGCAC	301
Qy	219	GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla	238
Db	302	CAGAACCAACCCCGGCTGTGCTGCCAGTAGCTCCCGATTGGAATCAITTCGCGCA	361
Qy	239	GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla	258
Db	362	GGGTCAACGACTCGGTGCTTACTCCACGCTGGGACACGAGCAGCAGAACCTCGCG	421
Qy	259	LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer	278
Db	422	CTGCATATCGTTGTTGGGACTCGCTGGCCATGATGTCTCCTCAGTCCACCAACACAGC	481
Qy	279	ThrLeuLeuArgTyrSerValSerLeuGlyTyrGlyPheTyrGlyAspIleIleIys	298
Db	482	ACACTCTTCTGCTACTCCGTGCTCCCTGCTGGGTACGGCTTCTACGGGGACATCATCAAG	541
Qy	299	AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr	318
Db	542	GACAGTGAGAAGAAACGGTGGTCTTGGGCTTTGCCAGATACGACTTTTCAGGTTTAAAGACC	601
Qy	319	PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal	338
Db	602	TTTCCTCTCCACCACTGCTATGAAGGGACAGTGTCTTCTTCCCTCCGACCAACACACG	661
Qy	339	GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerIys	358
Db	662	GGATCTCCAAGGATAGAAAGCCCTGCCGGGAGGATGCTTTGTTTCAGGCAAGCAAG	721
Qy	359	GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal	378
Db	722	CAGCAGCTGGAGGAGGAGCAGAGAAGACACTGTATGTTTGGAGCTGCGGAGGACGTG	781
Qy	379	GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys	398
Db	782	GAGGAGTGGCAAGTCGTCTGTGGGAGTTTCTGGCCATCAATGCACAAAACATGCTCCTGT	841
Qy	399	AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer	418
Db	842	GCTGTGTCGGGAGCCCCAGGGGCTCTCCCGGCTGCCACCTTGGGAGAGCGGCTCTTCT	901
Qy	419	AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis	438
Db	902	GACCTCATCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCAC	961
Qy	439	ThrAsnGlnGlnAspGln	444

```

Db      962 ACCAACCCAGCAGGACCAG 979
RESULT 15
AAS77731
ID   AAS77731 standard; cDNA; 1570 BP.
XX   AC
XX   AAS77731;
XX   13-FEB-2002 (first entry)
XX   DNA encoding novel human diagnostic protein #13535.
XX   Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX   Homo sapiens.
XX   WO200175067-A2.
XX   11-OCT-2001.
XX   30-MAR-2001; 2001WO-US008631.
XX   31-MAR-2000; 2000US-00540217.
XX   23-AUG-2000; 2000US-00649167.
XX   (HYSE-) HYSEQ INC.
XX   Drmanac RT, Liu C, Tang YT;
XX   WPI; 2001-639362/73.
XX   P-PSDB; ABG13544.
XX   New isolated polynucleotide and encoded polypeptides, useful in
XX   diagnostics, forensics, gene mapping, identification of mutations
XX   responsible for genetic disorders or other traits and to assess
XX   biodiversity.
XX   Claim 1; SEQ ID NO 13535; 103pp; English.
XX   The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX   sequences. (I) is useful as hybridisation probes, polymerase chain
XX   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX   and in recombinant production of (II). The polynucleotides are also used
XX   in diagnostics as expressed sequence tags for identifying expressed
XX   genes. (I) is useful in gene therapy techniques to restore normal
XX   activity of (II) or to treat disease states involving (II). (II) is
XX   useful for generating antibodies against it, detecting or quantitating a
XX   polypeptide in tissue, as molecular weight markers and as a food
XX   supplement. (II) and its binding partners are useful in medical imaging
XX   of sites expressing (II). (I) and (II) are useful for treating disorders
XX   involving aberrant protein expression or biological activity. The
XX   polypeptide and polynucleotide sequences have applications in
XX   diagnostics, forensics, gene mapping, identification of mutations
XX   responsible for genetic disorders or other traits to assess biodiversity
XX   and to produce other types of data and products dependent on DNA and
XX   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX   coding sequences of the invention. Note: The sequence data for this
XX   patent did not appear in the printed specification, but was obtained in
XX   electronic format directly from WIPO at
XX   ftp.wipo.int/pub/published_pct_sequences
XX   SQ   Sequence 1570 BP; 349 A; 446 C; 456 G; 319 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      4.16e-117      Length:      1570
Score:          1386.50        Matches:      278
Percent Similarity: 79.10%      Conservative: 21
Best Local Similarity: 73.54%    Mismatches:   45
Query Match:     48.01%         Indels:       34
DB:              5              Gaps:         6

```

```

US-10-631-958-10 (1-537) x AAS77731 (1-1570)
Qy      161 AlaserIleThrThrAspIleIleValThrGluHisAlaasnGlnAla-LysGluThrLe 180
      ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      437 AGCAGCATCCCGCACAGCAGCGTGTGTGTGGACATGCCACACATCCCGGGAGAAAAC 496
Qy      180 uTyrgluIleAsnIleAspIlystYrAspGlyIleValCysValGlyGlyAspGlyMetPh 200
      ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      497 ACACCTGACCCAGGTGGACAGA-----TGCTGTGGG----- 527
Qy      200 eSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp---Gl 219
      ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      528 -----AACATCTTGCAAGGGCGGAAGGCTGCTGCCCGAGGCTCGG 571
Qy      219 nAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGl 239
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      572 AAACACCCCGCTCCCGCCATCTGCCT-----GTGCAGTGGG 610
Qy      239 ySerThrAspCysValCysTyrSerThrValGlyThrSerAspAla--GluThrSerAla 258
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      611 CACGGCAGCTGTGCCTGGGCTGCCAGCTGTGGACGCATGCCCTCTGTGAGGCTCGAGG 670
Qy      259 LeuHisIleValVal-----GlyAspSerLeuAlaMetAspValSerSerValHis 275
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      671 CTTCACTCCAGGATGCAGAGCCCGGGGACTCGCTGGCCATGGATGTCTCTCAGTCCAC 730
Qy      276 HisAsnSerThrLeuLeuArgTyrSerValSerLeuGlyTyrGlyPheTyrGlyAsp 295
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      731 CACAAGACACACTCTTCCGTACTCCGTCTCTGGCTGTGGCTTCTACGGGGAC 790
Qy      296 IleIleAspSerGluIlystYsArgTyrLeuGlyLeuAlaArgTyrAspPheSerGly 315
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      791 ATCATCAAGGACAGTGAAGAAACGGTGTGGTCTTGCCAGATACGACTTTTCAGGT 850
Qy      316 LeuIlystPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGln 335
Db      851 TTAAGACCTTCTCTCCACCACTCTATGAAGGACAGTGTCTCTCTCCCTCGACAA 910
Qy      336 HisThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArg 355
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      911 CACACGGTGGGATCTCAAGGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTGCAGG 970
Qy      356 GlnSerIleGlnGlnLeuGluGluGlnLysIleAlaLeuTyrGlyLeuGluAlaAla 375
Db      971 CAAAGCAAGCAGCAGCTGGAGGAGGAGCAGAGAAAGACACTGTATGTTTGGAGCTGCG 1030
Qy      376 GluAspValGluGluTrpGlnValValCysGlyIlystPheLeuAlaIleAsnAlaThrAsn 395
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1031 GAGGACGTGGAGGAGTGGCAAGTCGTCTGTGGGAAGTTTCTGGCCCATCAATGCCACAAAC 1090
Qy      396 MetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAsp 415
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1091 ATGCTCTGTCTTGTGCGCGAGCCCGAGGGGCTCTCCCGGCTGCCCATTTGGGAGAC 1150
Qy      416 GlySerSerAspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeu 435
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1151 GGGTCTTCTGACCTCATCTCATCGGAATGCTCCAGGTTCAATTTCTGAGATTCTC 1210
Qy      436 IleArgHisThrAsnGlnGlnAspGlnPheAspPheThrPheValGluValTyrArgVal 455
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1211 ATCAGGCACACCAACGAGGAGGACCAAGTTGAGTTTCACTTTTGTGAAGTTTATCGGTC 1270
Qy      456 LysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGly 475
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1271 AAGAAATTCCAGTTTACGTGAAGCACATGGAGGATGAGGACAGCCACCTCAAGGAGGGG 1330
Qy      476 GlyLysLysArgPheGlyHisIleCysSerSerHisProSerCysCysThrValSer 495
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1331 GGAAGAGGCGCTTTGGGCACATTTTGAGCAGCAGCACCCTCTCTGCTGTGACCGTCTCC 1390
Qy      496 AsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGluValArgVal 515
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1391 AACAGCTCTTGGAACTGCGACGGGGAGGTCTCTGCACAGGCCCTGCCATCGAGGTGAGG--- 1447

```

Qy 516 HisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnPro 531  
Db 1448 -----GCCGTGGTGATGGCAAGAGGCTGCGAAAGGTCCCA 1483

Search completed: September 6, 2005, 11:25:27  
Job time : 893.053 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: September 5, 2005, 23:41:54 ; Search time 1512.64 Seconds  
(without alignments)  
2324.913 Million cell updates/sec  
Title: US-10-631-958-10  
Perfect score: 2888  
Sequence: 1 NGATGAEPLOSLVWVKQR.....QLVRLFARGIENPKPDHS 537  
Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0  
Searched: 7338684 seqs, 3274456166 residues  
Total number of hits satisfying chosen parameters: 14677368  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries  
Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US10631958/runat\_02092005\_165813\_4470/app.query.fasta\_1.1941  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOFCLO=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -WAIRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10631958 @CGN 1 1 1588 @runat\_02092005\_165813\_4470  
-NCPD=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100  
-LONGLOG -DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq.\*  
22: /cgn2\_6/ptodata/1/pubpna/US10J\_NEW\_PUB.seq.\*  
23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq.\*  
24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2888	100.0	1614	10 US-09-969-896-9	Sequence 9, Appli
2	2888	100.0	1614	19 US-10-631-958-9	Sequence 9, Appli
3	2888	100.0	1740	18 US-10-262-511-39	Sequence 39, Appl
4	2888	100.0	4413	10 US-09-969-896-16	Sequence 16, Appl
5	2888	100.0	4413	19 US-10-631-958-16	Sequence 16, Appl
6	2888	100.0	4429	19 US-10-618-941-55	Sequence 55, Appl
7	2881	99.8	4432	17 US-10-120-988-148	Sequence 148, App
8	2880	99.7	4463	16 US-10-315-597A-1	Sequence 1, Appli
9	2649.5	91.7	1840	9 US-09-784-810A-5	Sequence 5, Appli
10	2649.5	91.7	1840	22 US-10-876-281-5	Sequence 5, Appli
11	1640.5	56.8	979	10 US-09-969-896-1	Sequence 1, Appli
12	1640.5	56.8	979	19 US-10-631-958-1	Sequence 1, Appli
13	1032	35.7	817	19 US-10-115-635-247	Sequence 8, Appli
14	871	30.2	550	10 US-09-969-896-8	Sequence 8, Appli
15	871	30.2	550	19 US-10-631-958-8	Sequence 8, Appli
16	811	28.1	474	10 US-09-969-896-4	Sequence 4, Appli
17	811	28.1	474	19 US-10-631-958-4	Sequence 4, Appli
C 18	656	22.7	522	9 US-09-784-810A-7	Sequence 7, Appli
C 19	656	22.7	522	22 US-10-876-281-7	Sequence 7, Appli
20	616.5	21.3	2657	20 US-10-425-115-176413	Sequence 176413,
21	573.5	19.9	2079	18 US-10-425-114-34530	Sequence 34530, A
22	573.5	19.9	2079	20 US-10-425-115-86211	Sequence 86211, A
23	525	18.2	329	10 US-09-969-896-5	Sequence 5, Appli
24	525	18.2	329	19 US-10-631-958-5	Sequence 5, Appli
C 25	480.5	16.6	2189	19 US-10-437-963-92801	Sequence 92801, A
26	437.5	15.1	564	22 US-10-477-445-46	Sequence 46, Appl
27	399.5	13.8	1833	19 US-10-437-963-30650	Sequence 30650, A
28	389	13.5	2084	18 US-10-424-599-47396	Sequence 47396, A
29	376	13.0	2166	20 US-10-425-115-64890	Sequence 64890, A
30	371.5	12.9	1869	18 US-10-425-114-34806	Sequence 34806, A
31	371	12.8	339	9 US-09-783-590-5271	Sequence 5271, Ap
32	363	12.6	3362	20 US-10-425-115-29002	Sequence 29002, A
33	347	12.0	2014	18 US-10-424-599-59419	Sequence 59419, A
34	344.5	11.9	1512	20 US-10-425-115-64892	Sequence 64892, A
35	343.5	11.9	382	9 US-09-784-810A-9	Sequence 9, Appli
36	343.5	11.9	382	22 US-10-876-281-9	Sequence 34081, A
37	339	11.7	1774	18 US-10-425-114-34081	Sequence 95148, A
38	331.5	11.5	1386	19 US-10-437-963-95148	Sequence 77896, A
39	330.5	11.4	3008	19 US-10-437-963-77896	Sequence 32139, A
40	326.5	11.3	1857	20 US-10-425-114-32139	Sequence 97849, A
41	326.5	11.3	2247	18 US-10-425-115-97849	Sequence 58701, A
42	322	11.1	1916	19 US-10-437-963-58701	Sequence 25, Appl
43	319	11.0	2609	17 US-10-348-052-25	Sequence 25, Appl
44	319	11.0	2609	19 US-10-622-011-25	Sequence 79424, A
45	316.5	11.0	2284	20 US-10-425-115-79424	

ALIGNMENTS

RESULT 1  
US-09-969-896-9  
; Sequence 9, Application US/09969896  
; Publication No. US20030125533A1  
; GENERAL INFORMATION:  
; APPLICANT: Kossida, Sophia  
; TITLE OF INVENTION: Regulation of human Sphingosine  
; FILE OF INVENTION: Kinase-Like Protein  
; FILE REFERENCE: 004974.00594  
; CURRENT APPLICATION NUMBER: US/09/969,896  
; CURRENT FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: US 60/238,005  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: US 60/314,113  
; PRIOR FILING DATE: 2001-08-23  
; NUMBER OF SEQ ID NOS: 16

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1614
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-9

Alignment Scores:
Pred. No.: 0 Length: 1614
Score: 2888.00 Matches: 537
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-10-631-958-10 (1-537) x US-09-969-896-9 (1-1614)

Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnGlnArg 20
Db 1 ATGGGGCGACGGGGCGGGCGAGCCGCTGCAATCCGCTGCTGGTGAAGCAGCAGCGC 60

Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40
Db 61 TGGCCCGTGAACCTGGAGCCCGCGGGCTCTGCTGGCTGGTGGCGAGCCGGGGCCC 120

Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal 60
Db 121 GGACGGGGCCCCCGGGCGGGATGCCCTGCTCTGTGCTGTATCTGAGATCATCCCGTT 180

Qy 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80
Db 181 GAGGAAACAGACGTTTACGGGAAACATCAAGGCAGTGGAAATGGCAGAAATGGAAAG 240

Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrpLysTrpAla 100
Db 241 CCTTACGCTTTTACAGTTTCACTGTGTAAAGAGACGACGGCCCGCTGGAACTGGCG 300

Qy 101 GlnValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120
Db 301 CAGGTGACTTTCTGGTGTCCAGAGAGGAGCTGTGTCTCTGTGGTGCAGACCTCGGG 360

Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
Db 361 GAGATGCTGGAGAGCTGACGCTCCAGACCAAGCAATTTACTGTATTATCAACCCGTTT 420

Qy 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
Db 421 GGAGGAAAGGACCAAGGCAAGCGGATATATGAAAGAAAGTGGCACCCTGTTTACCTTA 480

Qy 161 AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180
Db 481 CCTTCCATCACCCTGACATCATCGTTACTGAAATGCTATATCAGGCCAAGGACTCTG 540

Qy 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200
Db 541 TATGAGATTAACATAGACAAATACAGCGCATCGTGTGTGGCGGAGATGATGTTTC 600

Qy 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220
Db 601 AGCGAGGTGCTGCGGGTCTGATTTGGGAGGACGACAGAGCGCGGGGTTCGACCAAG 660

Qy 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240
Db 661 CACCCCGGGCTGTGCTGGTCCCGAGTAGCTCCGATTTGAATTCATTTCCCGAGGGTCA 720

Qy 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
Db 721 ACGGACTCGGTGTGTTACTTCCACCGTGGGCACCGAGCAGCAAAACCTCGGCGCTGCAT 780

Qy 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 280
Db 781 ATCGTTTGTGGGACTCGCTGGCCATGATGTGCTCTCAGTCCACCAACAGCACACTC 840
```

```
Qy 281 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer 300
Db 841 CTTTGGCTACTCCGTGTCCTGCTGGGCTACGGCTTCTACGGGACATCATCAAGCAGT 900

Qy 301 GluLysLysArgTrpTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 320
Db 901 GAGAAGAAACGGTGGTGGTCTTGGCAGATAGACTTTTCAGGTTTAAAGACCTCTCCTC 960

Qy 321 SerHisHisCysTyrGlyGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340
Db 961 TCCACCACTGCTATGAAAGGACAGTGTCTTCTCTCTGCAACACACAGTGGGATCT 1020

Qy 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360
Db 1021 CCAAGGGATAGGAAGCCCTCGCGGAGGATGCTTTGTTTGCAGGCAAGCAAGCAGCAG 1080

Qy 361 LeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 380
Db 1081 CTGGAGGAGGAGCAGAAAGACACTGTATGTTTGGAAAGCTCGGAGGACGTGGAGGAG 1140

Qy 381 TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys 400
Db 1141 TGGCAAGTCTGCTGTGGAAAGTTTCTGCCCATCAATGCCCAACACATGTCTGTGCTGT 1200

Qy 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420
Db 1201 GCGCGAGGCCCGAGGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCTGACCTC 1260

Qy 421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440
Db 1261 ATCTCTATCCGAAATCTCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCACACCAAC 1320

Qy 441 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 460
Db 1321 CAGCAGACCAAGTTTGACTTCACTTTTGTGTAAGTTTATCGCTCAAGAAATTCAGTTT 1380

Qy 461 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe 480
Db 1381 ACGTGGAAGCACATGGAGGATGAGGACAGCGACTCAAGGAGGGGGGGAAGCGCTTT 1440

Qy 481 GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn 500
Db 1441 GGGCACATTTTGCAGACGCCACCCCTCTGCTGTGTCACCGCTCTCCAAACAGCTCTGGAAC 1500

Qy 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520
Db 1501 TGGACGGGGAGGTCTCTGCACGCCCTGCCATCGAGGTCAAGTCCACTGCCAGCTGGTT 1560

Qy 521 ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 537
Db 1561 CGACTCTTTGCAGGAGAAATTGAAGAGAAATCCGAAGCCAGACTCACACAGC 1611

RESULT 2
US-10-631-958-9
; Sequence 9, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631,958
; PRIOR FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1614
```

TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-631-958-9

## Alignment Scores:

Pred. No.: 0 Length: 1614  
Score: 2888.00 Matches: 537  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 19 Gaps: 0

US-10-631-958-10 (1-537) x US-10-631-958-9 (1-1614)

```
Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValValGlnGlnArg 20
Db 1 ATGGGGCGAGCGGGGCGCGGAGCGCGTCAATCCGTGCTGTGGGTGAAGCAGCAGCGC 60
Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40
Db 61 TGGCGCGTGAGCGCTGGAGCCGCGCGCGCTCTGCTGCGCTGGTGGCGAGCCCGGGGCC 120
Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleIleAlaVal 60
Db 121 GGAGCGCGCGCCCCCGCGGATGCCCTGCTCTGTCCTGTATCTGAGATCATGCCCGTT 180
Qy 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80
Db 181 GAGGAAACAGACGTTTACGGGAAACATCAAGGCGAGTGGGAAATGGCAGAAATGGAAAG 240
Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrpLysTrpAla 100
Db 241 CCTTACGCTTTTACAGTTTCACTGTGTAAAGAGAGCAGCAGCGCACCGCTGGAAGTGGCG 300
Qy 101 GlnValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120
Db 301 CAGGTGACTTCTTGTTGTCAGAGAGCAGCTGTGTCTACTTGTGCTGCAGACCCCTGCGG 360
Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
Db 361 GAGATGCTGGAGAACTCAGCTCCAGACCAAGCAATTTACTGGTATTATCAACCCGTTT 420
Qy 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
Db 421 GGAGAAAGGACCAAGGAGCGGATATGAAAGAAAGTGGCACCACTGTTTCACTTFA 480
Qy 161 AlaSerIleThrThrAspIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180
Db 481 GCCTCCATCACCTGACATCATCTGTTACTGAACATGCTTAATCAGGCCCAAGAGACTCTG 540
Qy 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyLysGlyMetPhe 200
Db 541 TATGAGATTACATAGACAAATACACGCGCATCGTCTGTGCGGGAGAGATGTTATGTT 600
Qy 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220
Db 601 AGCGAGGTGCTGCACGGTCTGATTGGAGGACGACAGAGCGCGCGGGTGCACCAAGAC 660
Qy 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240
Db 661 CACCCCGGGCTGTGCTGTCCTCCAGTAGCTCCCGATTGGAAATCATTTCCCGAGGGTCA 720
Qy 241 ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
Db 721 ACGGACTGGTGTGTTACTCCACCGTGGGACACGACGAGCAAACTCCGCGCTGCAT 780
Qy 261 IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu 280
Db 781 ATCGTGTGTTGGGACTCGCTGGCCATGATGTCTCTCAGTCCACCAACAGCAGCAGCTC 840
Qy 281 LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleAspSer 300
Db 841 CTTTCGTACTCCGTGCTGCTGGCTACCGCTTCTACGGGGACATCATCAAGGACAGT 900
```

```
Qy 301 GluLysLysArgTrpTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu 320
Db 901 GAGAAGAAACCGTGGTGGTCTTGGCCAGATACGACTTTTCAGGTTTAAAGACCTTCCTC 960
Qy 321 SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer 340
Db 961 TCCACCACTGCTATGAAGGAGCAGTGTCTTCTCTCCCTGCACACACACCGTGGGATCT 1020
Qy 341 ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln 360
Db 1021 CCAAGGATAGGAAGCCCTGCGGGCAGGATGCTTTGTTTGCAGGCAAGCAAGCAGCAG 1080
Qy 361 LeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu 380
Db 1081 CTGGAGGAGGAGCAGAAAGACACTGTATGTTTGGAAAGCTTGGAGAGCAGCTGGAGGAG 1140
Qy 381 TrpGlnValValCysGlyLysPheLeuAlaIleAlaThrAsnMetSerCysAlaCys 400
Db 1141 TGGCAAGTCGCTCTGGGAAGTTTCTGGCCATCAATGCCACAAACATGCTCTGTCTTGT 1200
Qy 401 ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu 420
Db 1201 GCGCGAGGCCCAAGGGCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCTGACCTC 1260
Qy 421 IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn 440
Db 1261 ATCCTCATCCGAAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCACACCAAC 1320
Qy 441 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 460
Db 1321 CAGCAGACCACTTGAATTTTCTTGTGTAAGTTTATCGCGTCAAGAAATTCAGTTT 1380
Qy 461 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyGlyLysLysArgPhe 480
Db 1381 ACGTGGAAGCACATGGAGGATGAGACAGCGACCTCAAGAGGGGGGGAAGAGCGCTTT 1440
Qy 481 GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn 500
Db 1441 GGGCACATTTGCAGCAGCCACCCCTCTCTGCTGTGTCACCGTCTCCAACAGCTCTTGAAC 1500
Qy 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520
Db 1501 TGGACCGGGAGGTCTTGCACAGCCCTGCTGCTGTCAGGTTCAGAGTCCACTGCCAGCTG 1560
Qy 521 ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 537
Db 1561 CGACTCTTTGCACAGGGAATTTGAAGAGAAATCCGAGGCCAGACTTCACAGC 1611
```

## RESULT 3

US-10-262-511-39

; Sequence 39, Application US/10262511

; Publication No. US20040038223A1

; GENERAL INFORMATION:

; APPLICANT: Smithson, Glenda

; APPLICANT: Millet, Isabelle

; APPLICANT: Peyman, John A.

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Ju, Jingfang

; APPLICANT: Li, Ji

; APPLICANT: Guo, Xiaojia (Sasha)

; APPLICANT: Patturajan, Meera

; APPLICANT: Spytek, Kimberly A.

; APPLICANT: Edinger, Shlomit R.

; APPLICANT: Ellerman, Karen

; APPLICANT: Malyankar, Uriel M.

; APPLICANT: Ort, Tatiana

; APPLICANT: Gorman, Linda

; APPLICANT: Zerkusen, Bryan D.

; APPLICANT: Anderson, David W.

; APPLICANT: Zhong, Mei

; APPLICANT: Catterton, Elina

; APPLICANT: Ji, Weizhen





Db 1396 CAGCAGGACGAGTTTGGACTTTCACATTTTGGTGAAGTTTATCCGGGTCAAGAAATCCAGTTT 1455  
 Qy 461 ThrSerLyHisMetGluAspGluAspSerAspLeuLysGluGlyGlyLysLysArgPhe 480  
 Db 1456 ACGTCGAAGCACATGGAGGATGAGACAGCGACCTCAAGGAGGGGGGAGAGAGCGCTTT 1515  
 Qy 481 GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn 500  
 Db 1516 GGGCACATTTGCAGCAGCCACCCCTCCTGCTGTGCTGCACCGCTCTCCAAACAGCTCCTCGAAC 1575  
 Qy 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520  
 Db 1576 TGCACCGGGGAGGTCTCTGCACAGCCCTGCCATCGAGGTCAAGTCCACTGCCAGCTGGTT 1635  
 Qy 521 ArgLeuPheAlaArgGlyIleGluGluAsnProLysProAspSerHisSer 537  
 Db 1636 CGACTCTTTGCACGAGGAATGAAGAAATCCGAGCCAGCTCACACAGC 1686  
 RESULT 4  
 US-09-969-896-16  
 ; Sequence 16, Application US/09969896  
 ; Publication No. US20030125533A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Koseida, Sophia  
 ; TITLE OF INVENTION: Regulation of human Sphingosine  
 ; TITLE OF INVENTION: Kinase-Like Protein  
 ; FILE REFERENCE: 004974.00594  
 ; CURRENT APPLICATION NUMBER: US/09/969,896  
 ; CURRENT FILING DATE: 2001-10-04  
 ; PRIOR APPLICATION NUMBER: US 60/238,005  
 ; PRIOR FILING DATE: 2000-10-06  
 ; PRIOR APPLICATION NUMBER: US 60/314,113  
 ; PRIOR FILING DATE: 2001-08-23  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 16  
 ; LENGTH: 4413  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-969-896-16  
 Alignment Scores:  
 Pred. No.: 0 Length: 4413  
 Score: 2888.00 Matches: 537  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0  
 US-10-631-958-10 (1-537) x US-09-969-896-16 (1-4413)  
 Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnArg 20  
 Db 76 ATGGGGCGCAGCGGGGGCGGAGCCGCTGCATTCGGTGTGTGGGTGAAGCAGCAGCGC 135  
 Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40  
 Db 136 TGCGCCGTGAGCCTGGAGCCCGCGCGGCTCTGCTGGCTGTGTGGCGAGCCCGGGGCC 195  
 Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleIleAlaVal 60  
 Db 196 GGAGCGCGCGCCCGCGCGCGATGCCCTGCTGTGCTGTATCTGAGATCATCGCGCTT 255  
 Qy 61 GluGluThrAspValHisGlyLysHisGlnGlySerClyLysTrpGlnLysMetGluLys 80  
 Db 256 GAGGAAACACACGCTTCACGGGAAACATCAAGGCAGTGGAAAAATGGCAGAAATGGAAG 315  
 Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTrpLysTrpAla 100  
 Db 316 CCTTACGCTTTACAGTTTCACTGTGTAAAGAGACGACGCGCACCGCTGGAAAGTGGCG 375  
 Qy 101 GlnValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120

376	Db	CAGGTGACTTTCTGGTGTCCAGAGGACAGCTGTGTCACTTGTGGCTGCGACAGCCCTCGGG	433
121	Qy	GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe	140
436	Db	GAGATGCTGGAGAAAGCTGACGTCCAGACCAGACATTTACTGGTATTATTATCAACCCGTTT	495
141	Qy	GlyGlyGlyGlnGlnGlyArgIleTyrGluArgLysValAlaProLeuPheThrLeu	160
496	Db	GGAGGAAAAGGACAAGCGGATATATGAAGAAAAGTGGCACCACTGTTCACTTCA	555
161	Qy	AlaSerIleThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu	180
556	Db	GCCTCATCACCACTGACATCATCTGTTACTGAACATGCTAATCAGGCCAAGGAGACTCTG	615
181	Qy	TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe	200
616	Db	TATGAGATTACATAGACAATACGACGGCATCGTCTGTGTCGGCGAGATGTTATGTTCT	675
201	Qy	SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn	221
676	Db	AGCGAGGTGTCGACGGTCTGATTGGAGGACGACGAGGAGCGCGGGTTCGACCAGAAC	735
221	Qy	HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer	240
736	Db	CACCCCGGGCTGTGTGGTCCCAGTAGCTCCGGATTGGAATCATCTCCGACGGGTCA	795
241	Qy	ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis	260
796	Db	ACGGACTGCGTGTGTTATCTCCACGCTGGGCACGACGACGAGAAACCTCGGGCTCGAT	855
261	Qy	IleValValGlyAspSerLeuAlaMetAspValSerSerValHisAsnSerThrLeu	280
856	Db	ATCGTTGTTGGGACTCGCTGGCCATGGATGTGCTCTCAGTCCACCACAACAGCACACTC	915
281	Qy	LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer	300
916	Db	CTTCGTACTCCGTGCTCCTGTCTGGGCTACGGCTTCTACGGGACATCATCAAGGACGT	975
301	Qy	GluLysIleArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu	320
976	Db	GAGAGAACCGGTGGTGGTCTGTGCCAGATACGACTTTTCAGGTTTAAAGACCTTCCTC	1035
321	Qy	SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer	340
1036	Db	TCCCACCCTGCTATGAAGGACAGTGCCTCTCTCCCTGTCACAACACACGSGTGGATCT	1095
341	Qy	ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln	360
1096	Db	CCAAGGGATAGGAAGCCCTCCCGGGAGGATGCTTTGTTTTCAGAGGCAAAAGCAAGCAG	1155
361	Qy	LeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu	380
1156	Db	CTGGAGGAGAGCAGAAGAAAGACACTGTATGTGTTTGGAGCTCGGAGGACGTGGAGGAG	1215
381	Qy	TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys	400
1216	Db	TGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCACAACATGTCTCTGTGCTGT	1275
401	Qy	ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu	420
1276	Db	CGCCGGAGCCCCCAGGGGCGCTCTCCCGGGCTGCCACTTGGGAGACGGGTCTTCTGACCTC	1335
421	Qy	IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn	440
1336	Db	ATCCTCATCCGGAATATGCTCCAGGTTCAATTTTCTGAGATTCTCATCAGGCACACCAAC	1395
441	Qy	GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe	460
1396	Db	CAGCAGGACGATTTGACTTTCATTTTGTGAAGTTTATCGCGTCAAGAANAATCCAGTTT	1455
461	Qy	ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe	480
1456	Db	ACGTCTGAAGCAGATGAGGATGAGGACGACCTCAAGAGGGGGGAAGAAGCGCTTT	1515



Db 1516 GGGCACANTTTCAGACAGCCACCCCTCTCTGTGCTGCACCGGTCTCCAAACAGCTCTCTGGAAC 1575  
 Qy 501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520  
 Db 1576 TGCAGCGGGAGAGTCTCTGCACAGCCCTGCCATCCAGGTCCAGATCCACTGCCAGTGTGTT 1635  
 Qy 521 ArgLeuPheAlaArgGlyTleGluGluAsnProLysProIlePheSerHisSer 537  
 Db 1636 CGACTCTTTGCACGAGGAAATTGAAGAGAATCCGAAGCCAGACTCACACAGC 1686  
 RESULT 6  
 US-10-618-941-55  
 ; Sequence 55, Application US/10618941  
 ; Publication No. US2004019792A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WHYTE, DAVID  
 ; APPLICANT: MANNING, GERARD  
 ; APPLICANT: CAENEPEEL, SEAN  
 ; TITLE OF INVENTION: NOVEL KINASES  
 ; FILE OF INVENTION: 034536-0321  
 ; CURRENT APPLICATION NUMBER: US/10/618,941  
 ; CURRENT FILING DATE: 2003-07-15  
 ; PRIOR APPLICATION NUMBER: 60/395,632  
 ; PRIOR FILING DATE: 2002-07-15  
 ; NUMBER OF SEQ ID NOS: 143  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 55  
 ; LENGTH: 4429  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-618-941-55  
 Alignment Scores:  
 Pred. No.: 0 Length: 4429  
 Score: 2888.00 Matches: 537  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 19 Gaps: 0  
 US-10-631-958-10 (1-537) x US-10-618-941-55 (1-4429)  
 Qy 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTrpValLysGlnArg 20  
 Db 92 ATGGGGCGAGCGGGGGCGCGAGCCGCTGCAATCCGTCTGTGGGTGAAGCAGCAGCGC 151  
 Qy 21 CysAlaValSerLeuGluProAlaAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40  
 Db 152 TGGCGCCGTGAGCCTGAGCCCGCGCGGGCTCTGCTGGCGCTGGTGGCGAGCCCGGGGGCCC 211  
 Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleIleAlaVal 60  
 Db 212 GGAAGCGGGCGCCCCCGGGCGGATGCTCTGTGCTGTATCTGAGATCATCGCGGTT 271  
 Qy 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80  
 Db 272 GAGGAACACAGCGTTTACGGGGAACATCAGGCAGTCGAAATGGCAGAAATGGAAG 331  
 Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAla 100  
 Db 332 CCTTACGCTTTTACAGTTCCTACTGTGTAAAGAGACAGCAGCGCACCCCTCGAAGTGGCG 391  
 Qy 101 GlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120  
 Db 392 CAGGTGACTTTCGTGTTCAGAGAGACAGCTGTGTCTACTGTGTGCTGTCAGACCCCTGCGG 451  
 Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140  
 Db 452 GAGATGCTGGAGAGCTGACGTCAGACCAAGCAATTTACTGGTATTTATCAACCCGTTT 511  
 Qy 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160  
 Db 512 GGAGAAAGAGCAAGCAAGCGGATATATGAAGAAAGATGGCACCACCTGTCACCTTA 571

Qy	521	ArgLeuPheAlaArGlylleGlUgLuAsnProLysProAspSerHisSer	537
		{     }	
Db	1652	CGACTCTTTGCACGAGGAATTGAAGAGAATCCGAAGCCAGACTCACACAGC	1702
 RESULT 7			
US-10-988-148			
; Sequence 148, Application US/10120988			
; Publication No. US20030219745A1			
; GENERAL INFORMATION:			
; APPLICANT: Tang, Y. Tom			
; APPLICANT: Goodrich, Ryle			
; APPLICANT: Liu, Chenghua			
; APPLICANT: Ren, Feiyang			
; APPLICANT: Wang, Dunrui			
; APPLICANT: Dmanac, Radjoje T.			
; TITLE OF INVENTION: No. US20030219745A1el Nucleic Acids and			
; FILE REFERENCE: 802CON			
; CURRENT APPLICATION NUMBER: US/10/120,988			
; CURRENT FILING DATE: 2002-04-11			
; PRIOR APPLICATION NUMBER: 09/774,528			
; PRIOR FILING DATE: 2001-01-30			
; NUMBER OF SEQ ID NOS: 441			
; SOFTWARE: pc_FL_genes Version 2.0			
; SEQ ID NO 148			
; LENGTH: 4432			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (1129) .. (2817)			
US-10-120-988-148			
 Alignment Scores:			
Pred. No.:                  0                  Length:              4432			
Score:                     2881.00               Matches:              536			
Percent Similarity:       99.81%            Conservative:        0			
Best Local Similarity:    99.81%            Mismatches:        1			
Query Match:              99.76%            Indels:              0			
DB:                         17                  Gaps:                0			
 US-10-631-958-10 (1-537) x US-10-120-988-148 (1-4432)			
Qy	1	MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValleuTrpVallysGlnArg	20
Db	1204	ATGGGGCGACGGGGCGCGAGCCCGCTGCAATCCGTGCTGTGGTGAAAGCAGACAGCGC	1263
Qy	21	CysAlaValSerLeuGluProAlaArgAlaLeuArgTrpTrpArgSerProGlyPro	40
Db	1264	TGCGCCGTGAGCCTTGAGCCCGCGCGGCTCTGTGCGCTGGTCGGCGAGGCCGGGGCCC	1323
Qy	41	GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal	60
Db	1324	GGAGCCGGCGCCCCCGCGCGGATGCCTGCTGTGGCTGTATCTGAGATCATCGCCGTT	1383
Qy	61	GluGluThrAspValHisGlyLysHissGlnGlySerGlyLysTrpGlnLysMetGluLys	80
Db	1384	GAGGAACACAGAGTTTCCACGGGAACATCAGCGCAGTGGAAAATTGGCAGAAAATCGAAAG	1443
Qy	81	ProTyrrAlaPheThrValHisCysVallysArgAlaArgHisaGrTrpLysTrpAla	100
Db	1444	CCTTAGCGTTTTACAGTTTCACTGTGTAAAGAGACACAGCGACCCTGGAAGTGGCGC	1503
Qy	101	GlnValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg	120
Db	1504	CAGGTGACTTTCTGGTGTCACAGAGGAGCAGCTGTGTCACTTTGTGGCTGCAGACCCTGGG	1563
Qy	121	GluMetLeuGluLysLeuThr-SerArgProLysHisLeuLeuValPheIleAsnPropHe	140
Db	1564	GAGATGCTTGGAGAAAGCTGACGTCAGACCACAAAGCAATTACTGTTATTATCAACCGTTT	1623
Qy	141	GlyGlyLysGlyGlnGlyLysArgiletyrGluArgLysValAlaProLeuPheThrLeu	160

Qy	161	AlaSerIleThrThrAspLlleValThrGluHisAlaAenGlnAlaLysGluThrLeu	180
Db	604	GCCTCCATCACCACGTGATCATCGTTACTGAAATGTCATAGGCCAACGGAGACTCTG	663
Qy	181	TyrGluIleAsnIleAspLysTyrAspGlyLeValCysValGlyGlyAspGlyMetPhe	200
Db	664	TATGAGATTTAACATAGACAATAACGAGGCATCGTGTCGGCGGAGATGGTATGTTCT	723
Qy	201	SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn	220
Db	724	AGCCAGGTGCTCAACGCTGATTTGGGAGGACGACAGAGGCGCGGGTCCGACAGAAC	783
Qy	221	HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer	240
Db	784	CACCCCGGGTGTCGTGCTCCCAGTAGCCTCCGATTGGAAATCATTTCCCGCAGGTCCA	843
Qy	241	ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis	260
Db	844	ACGACATCGGTGTGTACTCCACGCTGGGCACACGACGACGAGAATCCTCGCGCTGCAT	903
Qy	261	IleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSerThrLeu	280
Db	904	ATCGTGTGGGACTCGCTGGCCATGGATGTGTCTCAGTCCACCAACACAGCACACTC	963
Qy	281	LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspLlleLysAspSer	300
Db	964	CTTCGCTACTCCGTGCTCCCTGCTGGGCTACGGCTTCTACCGGGACATCATCAAGGACGT	1023
Qy	301	GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu	320
Db	1024	GAGAAGAACCGTGGTGGGTCTTGGCAGATACGACTTTTCAGGTTTAAAGACCTTCCTC	1083
Qy	321	SerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer	340
Db	1084	TCCACCACTGCTATGAAGGGACAGTGTCTTCTCTCCCTGCACNAACACACGCTGGATCT	1143
Qy	341	ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln	360
Db	1144	CCAAGGATAGGAAGCCCTCGCGGCAGGATGCTTTGTTTGCAGGCAAAGCAAGCAGCAG	1203
Qy	361	LeuGluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu	380
Db	1204	CTGAGAGAGAGCAGAAAGACACTATATGTTTGGANGCTCGGAGSACGTGGAGGAG	1263
Qy	381	TrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys	400
Db	1264	TGGCAGTGTCTGTGGGAAGTCTTGCCCATCATGCCCAAAACATGTCTCTGTGCTGTGT	1323
Qy	401	ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu	420
Db	1324	CGCCGAGCCCCAGGGCCCTCTCCCGGCTGCCACTTGGGAGACGGGTCTTCTGACCTC	1383
Qy	421	IleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHisThrAsn	440
Db	1384	ATCCTCATCCGAAAATGCTCCAAGTTCATATTTCTTGAGATTTCTCATCGCACACCAAC	1443
Qy	441	GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe	460
Db	1444	CAGCAGACCAAGTTTGACTTCACTTTTGTGTGAAGTTTATCCGCTCAAGAAATTCAGTTT	1503
Qy	461	ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe	480
Db	1504	ACGTGCAAGCACATGGAGGATGAGGACAGCGACTCAAGGAGGGGGGGAAGAGCGCTTT	1563
Qy	481	GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn	500
Db	1564	GGGCACATTTGCAGACACCAACCCCTCTGCTGTGTGACCGCTCTCCAACAGCTCTCTGAAC	1623
Qy	501	CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal	520
Db	1624	TGCGATGGGAGGTCTTGACACACCCCTGCCATCGAGTCAAGTCACTGCCAGCTGGTT	1683
Qy	521	ArgLeuPheAlaArgGlyLlleGluGluAsnProLysProAspSerHisSer	537



Db 1684 CGACTCTTTGACGAGGAATTAAGAGAAATCCGAAGCCAGACTCACAGC 1734  
RESULT 9  
US-09-784-810A-5  
; Sequence 5, Application US/09784810A  
; Patent No. US20020082203A1  
; GENERAL INFORMATION:  
; APPLICANT: RASTELLI, LUCA  
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING  
; TITLE OF INVENTION: SAME  
; FILE REFERENCE: 10716-08  
; CURRENT APPLICATION NUMBER: US/09/784.810A  
; CURRENT FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: 60/182,360  
; PRIOR FILING DATE: 2000-02-14  
; PRIOR APPLICATION NUMBER: 60/191,261  
; PRIOR FILING DATE: 2000-03-22  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 1840  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-784-810A-5  
Alignment Scores:  
Pred. No.: 8,37e-303 Length: 1840  
Score: 2649.50 Matches: 495  
Percent Similarity: 97.83% Conservative: 1  
Best Local Similarity: 97.63% Mismatches: 0  
Query Match: 91.74% Indels: 11  
DB: 9 Gaps: 1  
US-10-631-958-10 (1-537) x US-09-784-810A-5 (1-1840)  
Qy 42 AlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaValGlu 61  
Db 1 GCCGGCGCCCGCGCGGATGCCTCTCTGCTGTATCTGAGATCATCGCGTTGAG 60  
Qy 62 GluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLysPro 81  
Db 61 GAAACAGACGTTCAACGGGAAACATCAAGCGCAGTGGAAAAATGGCAAAATGGAAAGCCT 120  
Qy 82 TyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAlaGln 101  
Db 121 TACGCTTTTACAGTTCACTGTGTAAAGAGACACGCGCACCCGCTGGAAGTGGCGGAG 180  
Qy 102 ValThrPheTrpCysProGluGluGlnLeuLysHisLeuTrpLeuGlnThrLeuArgGlu 121  
Db 181 GTGACTTTCTGGTGTCCAGAGGAGCAGCTGTGTCTACCTTGTGCTGCAGACCTCGCGGAG 240  
Qy 122 MetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPheGly 141  
Db 241 ATGCTGGAGAAAGCTGACGCTCCAGACCAAGCAATTTACTGGTATTTATCAACCCGTTTGA 300  
Qy 142 GlyLysGlyGlnGlyLysArgIleTrpGluArgLysValAlaProLeuPheThrLeuAla 161  
Db 301 GGAAGAGACAAGCGACCGGATATATGAAAGAAAGTGGCACCACCTGTTCCACTTTAGCC 360  
Qy 162 SerIleThrThrAspIleIle-----ValThr 170  
Db 361 TCCATCACCACCTGACATCATCGGTAAACAAATTCATGTATGTTAACTATGTAGAAAGTAATTACT 420  
Qy 171 GluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTrpAspGly 190  
Db 421 GAACATGCTAATACAGGCCAAGGAGACTCTGTATGAGATTAAATAGACAAATACGACGCG 480  
Qy 191 IleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLysGlyValArg 210  
Db 481 ATCGTCTGTGCGCGGAGATGGTATGTTTACGAGAGGTGCTGCACGCTCTGATTTGGAGG 540  
Qy 211 ThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSer 230

Db 541 ACGCAGAGGAGCGCGCGGTCGACAGAACACCCCGGGCTGTGCTGCTCCCGTAGC 600  
Qy 231 LeuArgIleGlyIleIleProAlaGlySerThrAspCysValCysTrpSerThrValGly 250  
Db 601 CTCGGATTGGAATCATTTCCCGAGGTCACAGACTGCGTGTGTATCTACCCGCTGGC 660  
Qy 251 ThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAsp 270  
Db 661 ACCAGCGACGAGAAACCTCGGCGCTGCATATCGTTGTGGGACTCGCTGGCCATGGAT 720  
Qy 271 ValSerSerValHisHisAsnSerThrLeuLeuArgTrpSerValSerLeuGlyTyr 290  
Db 721 GTGTCTCTAGTCCACCAACACAGCACACTCTCTCGCTACTCCGTGTCCCTGCTGGGCTAC 780  
Qy 291 GlyPheTrpGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArg 310  
Db 781 GGCTTCTACCGGGACATCAAGGACAGTGAAGAAACCGTGGTGGGTCTTGGCCAGA 840  
Qy 311 TyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTrpGluGlyThrValSer 330  
Db 841 TACGACTTTTACGTTTAAAGACTTCTCTCCACCACTGCTATGAAGGACAGTGTCC 900  
Qy 331 PheLeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGly 350  
Db 901 TTCTCTCCCTGCACAACACACGCTGGGATCTCCAAGGATAGGAAGCCCTGCGGGCAGGA 960  
Qy 351 CysPheValCysArgGlnSerLysGlnLeuGluGlnGluLysLysAlaLeuTyr 370  
Db 961 TGCTTTGTTTGGAGGCAAGCAAGCAGCAGCTGGAGAGGAGGAGAGAAAGCACTGTAT 1020  
Qy 371 GlyLeuGluAlaAlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAla 390  
Db 1021 GGTTTGGAAGCTCGGAGGACGTGGAGAGTGGCAACTCGTCTGTGGGAAGTTTTCGCC 1080  
Qy 391 IleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAla 410  
Db 1081 ATCAATGCCACAAACATGTCTGTGCTGTGCGCGAGCCCGAGGGGCTCTCCCGGCT 1140  
Qy 411 AlaHisLeuGlyAspGlySerSerAspLeuIleLeuLysCysSerArgPheAsn 430  
Db 1141 GCCCAGCTGGGAGACGGGTCTTCTGACCTATCTCCGGAATGCTCCAGGTTCAAT 1200  
Qy 431 PheLeuArgPheLeuIleArgHisThrAsnGlnAspGlnPheAspPheThrPheVal 450  
Db 1201 TTTCTGAGATTCTCATCAGGCACACCAACACAGCAGACCACTTTGACTTCACTTTGTT 1260  
Qy 451 GluValTrpArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSer 470  
Db 1261 GAAGTTTATCGCGTCAAGAAATTCAGTTTACGTTCGAAGCACATGGAGGATGAGGACAGC 1320  
Qy 471 AspLeuLysGluGlyLysLysArgPheGlyHisIleCysSerSerHisProSerCys 490  
Db 1321 GACCTCAAGAGGGGGGAGAGAGCGCTTTGGGCACATTTGACAGACCCCTCTCTGC 1380  
Qy 491 CysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAla 510  
Db 1381 TGCTGACCGCTCTCCAACAGCTCTTGGAACTGCGAGCGGGAGGTCCTGACAGCCCTGCC 1440  
Qy 511 IleGluValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsn 530  
Db 1441 ATCGAGGTCAGGGTCCACTGCCAGCTGTTGCACTCTTTCACGAGGAATTTGAAGAAAT 1500  
Qy 531 ProLysProAspSerHisSer 537  
Db 1501 CCGAAGCCAGACTCACACAGC 1521  
RESULT 10  
US-10-876-281-5  
; Sequence 5, Application US/10876281  
; Publication No. US20050123942A1  
; GENERAL INFORMATION:  
; APPLICANT: RASTELLI, LUCA



; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/10/876,281
; PRIORITY FILING DATE: 2004-06-24
; PRIOR APPLICATION NUMBER: US/09/784,810
; PRIORITY FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIORITY FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIORITY FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1840
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-876-281-5

Alignment Scores:
Pred. No.: 8,376-303 Length: 1840
Score: 2649.50 Matches: 495
Percent Similarity: 97.83% Conservative: 1
Best Local Similarity: 97.63% Mismatches: 0
Query Match: 91.74% Indels: 11
Dbs: 22 Gaps: 1

US-10-631-958-10 (1-537) x US-10-876-281-5 (1-1840)

Qy 42 AlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaValGlu 61
Db 1 GCCCGCGCCCGCGCGGATGCCCTGCTGTGCTGTATCGAGATCATGCCGTTGAG 60

Qy 62 GluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLysPro 81
Db 61 GAAACAGAGCTTACGCGGAAACATCAAGCAGCTGGAATGCGAATAATGGAAGCCT 120

Qy 82 TyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAlaGln 101
Db 121 TACGCTTTTACAGTTTCACTGTGTAAAGAGAGCAGCAGCGCACCGCTGGAAAGTGGCGCAG 180

Qy 102 ValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArgGlu 121
Db 181 GTGACTTTTGTGTGTCAGAGAGAGCAGCTGTGTCTGCTGCTGCACACCTTGGCGGAG 240

Qy 122 MetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPheGly 141
Db 241 ATGCTGGAGAGCTACGCTCCAGACCAAGCATTTACTGGTATTTATCAACCCGTTTGA 300

Qy 142 GlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeuAla 161
Db 301 GGAAAAGGACAAGGCAAGCGGATATATGAAGAAAAGTGGCACCACTGTTACCTTAGCC 360

Qy 162 SerIleThrThrAspIleIle-----ValThr 170
Db 361 TCCATCACCACTGACATCATCGTAAACAAATTTCTATGTTAACTATGTAGAAGTAATTACT 420

Qy 171 GluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGly 190
Db 421 GAACATGTATATCAGCCAGAGAGACTCTGTATGAGATTAAATAGACAAATACGACGCG 480

Qy 191 IleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArg 210
Db 481 ATCGCTGTGTGCGCGGAGATGTTATGTTTCAGCGAGGTGCTGCACGGCTCTGATGGGAGG 540

Qy 211 ThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSer 230
Db 541 ACGCAGAGAGCGCGCGGTCGACAGAACCAACCCCGGCTGTGCTGCTCCCAAGTAGC 600

Qy 231 LeuArgIleGlyIleIleProAlaGlySerThrAspCysValCysTyrSerThrValGly 250
Db 601 CTCCGGATTGGATCATTTCCCGCAGGGTCAACGGAGCTGCGTGTGTACTTCCACCGTGGGC 660

Qy 251 ThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAsp 270
Db 661 ACCAGCGACGAGAAACCTCGCGCTGCATATCGTTGTGGGACTTCGCTGGCCATGGAT 720

Qy 271 ValSerSerValHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyr 290
Db 721 GTGTCTCAGTCCACCAACACAGCACACTCTCTGGTACTCGGTGCTCCTGCTGGGCTAC 780

Qy 291 GlyPheTyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArg 310
Db 781 GGCTTCTACGGGACATCATCAAGCAGTGAAGAAACGGTGGTGGTCTTGGCTTGGCCAGA 840

Qy 311 TyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSer 330
Db 841 TACGACTTTTTCAGGTTTAAAGACCTTCTCTCCCACTCTATGAAGGACAGTGTGTC 900

Qy 331 PheLeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGly 350
Db 901 TTCTCTCCCTGCACACACACCGTGGATCTCCAGGGNTAGNAGCCCTGCGGGCAGGA 960

Qy 351 CysPheValCysArgGlnSerLysGlnGlnLeuGluGlnLysLysAlaLeuTyr 370
Db 961 TGCCTTTTTCAGCGCAAAAGCAAGCAGCAGCTGGAGGAGGAGCAGAAAGCACTGTAT 1020

Qy 371 GlyLeuGluAlaAlaGluAspValGluGluTrpGlnValValCysGlyLysPheLeuAla 390
Db 1021 GGTGTGGAAGCTGCGGAGGACGTGGAGAGTGGCAAGTCTGTGTGGGAAGTTTCTGGCC 1080

Qy 391 IleAsnAlaThrAsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAla 410
Db 1081 ATCAATGCCACAAACATGTCTGTGTGTCGCGAGCCCCAGGGGCTCTTCCCGGCT 1140

Qy 411 AlaHisLeuGlyAspGlySerSerAspLeuIleLeuLysArgLysCysSerArgPheAsn 430
Db 1141 GCCCAGCTGGGAGACGGCTCTTCTGACCTCATCTCCGGAATGCTCCAGGTTCAAT 1200

Qy 431 PheLeuArgPheLeuIleArgHisThrAsnGlnGlnAspGlnPheAspPheThrVal 450
Db 1201 TTTCTGAGATTCTCATCAGGCACACCAACAGCAGGACCACTTTGACTTTCATCTTTGTT 1260

Qy 451 GluValTyrArgValLysPheGlnPheThrSerLysHisMetGluAspGluAspSer 470
Db 1261 GAAGTTTATCCGTCACAGAAATTCAGTTTACGTCGAGCAACATGGAGGATGAGGACAGC 1320

Qy 471 AspLeuLysGluGlyGlyLysArgPheGlyHisIleCysSerSerHisProSerCys 490
Db 1321 GACCTCAAGGAGGCGGGGAGAGAGCGCTTTGGGCACATTTGCAGCAGCCACCCCTCTGTC 1380

Qy 491 CysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAla 510
Db 1381 TGCTGCACCGTCTCCAAACAGCTCCTGGAACTGCGACGGGAGGTCTCTGCACAGCCCTGCC 1440

Qy 511 IleGluValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGluLeuAsn 530
Db 1441 ATCAGGTCAGGGTCCACTGCCAGCTGGTTCGACTCTTTGCACGAGGAATTTGAAGAGAAT 1500

Qy 531 ProLysProAspSerHisSer 537
Db 1501 CCGAAGCCAGACTCACACAGC 1521

RESULT 11
US-09-969-896-1
; Sequence 1, Application US/09969896
; Publication No. US20030125533A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06

```
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 979
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-1

Alignment Scores:
Pred. No.: 1.27e-183 Length: 979
Score: 1640.50 Matches: 314
Percent Similarity: 96.63% Conservative: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 56.80% Indels: 11
DB: 10 Gaps: 1

US-10-631-958-10 (1-537) x US-09-969-896-1 (1-979)

Qy 130 ProlysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgile 149
Db 2 CCAAAGCATTTACTGGTATTATCAACCGTTTGGAGAAAGGACAAAGCGCGGATA 61

Qy 150 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleile--- 168
Db 62 TATGAAAGAAAGTGGCACCACTGTTACCTTAGCCTCCATCACCCTGACATCATCGGT 121

Qy 169 -----ValThrGluHisAlaAsnGlnAlaLysGlu 178
Db 122 AACAAATTTCTATGTTAACTATGTAGAAAGTAAATTAATCTGAACATGCTAATCAGGCCAAGGAG 181

Qy 179 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 198
Db 182 ACTCTGTATGAGATTACATAGACAAATACGACGGCATCGTCTGTGTGCGGGAGATGGT 241

Qy 199 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValasp 218
Db 242 ATGTTACGAGGAGTGTGCACGGTCTGATTTGGGAGGACGACAGGAGCGCGGGTCCAC 301

Qy 219 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAla 238
Db 302 CAGAACACCCCGGGCTGCTGGTCCCAAGTAGCTCCGATTCGGATTCATTTCCGCGCA 361

Qy 239 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 258
Db 362 GGGTCAACCGACTCGGTGTGTACTTCCACCGTGGGACCCAGCAGCAGAAACCTCGCG 421

Qy 259 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 278
Db 422 CTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTGCTCTCAGTCCACCAACACAGC 481

Qy 279 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleileys 298
Db 482 ACACCTCTCTCTACTCCGTGCTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAG 541

Qy 299 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 318
Db 542 GACAGTGAGAAAGAACGGTGGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACC 601

Qy 319 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 338
Db 602 TTCTCTCTCCACCACTGCTATGAAAGGACAGTGTCTTCTCTCTCCCTGCACACACACAGGTG 661

Qy 339 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerIys 358
Db 662 GGATCTCCAAAGGATAGGAAGCCCTGCGGCGAGGATGCTTTGTTTTCAGGCAAGCAAG 721

Qy 359 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 378
Db 722 CAGCAGCTGGAGGAGGACAGAGAACATCTGTATGGTTTGGAGCTGCGGAGGAGCGTG 781

Qy 379 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaIleThrAsnMetSerCys 398
```

---

```
Db 782 GAGGAGTGGCAAGTCTGCTGTGGGAAGTTTCTGCCATCAATGCCACAAACATGCTCTGT 841
Qy 399 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 418
Db 842 GCTGTGTCGCGAGAGCCCGCAGGGGCTCTCCCGGCTGCCACCTTGGGAGACGGGTCCTCT 901
Qy 419 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 438
Db 902 GACCTCATCTCTCATCCGGAATGCTCCAGGTTCAATTTCTGAGATTTCTCATCAGGCAC 961

Qy 439 ThrAsnGlnGlnAspGln 444
Db 962 ACCAACCCAGCAGGACCAG 979

RESULT 12
US-10-631-958-1
; Sequence 1, Application US/10631958
; Publication No. US20040192580A1
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; TITLE OF INVENTION: Kinase-Like Protein
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/10/631.958
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US/09/969,896
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 979
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-958-1

Alignment Scores:
Pred. No.: 1.27e-183 Length: 979
Score: 1640.50 Matches: 314
Percent Similarity: 96.63% Conservative: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 56.80% Indels: 11
DB: 19 Gaps: 1

US-10-631-958-10 (1-537) x US-10-631-958-1 (1-979)

Qy 130 ProlysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArgile 149
Db 2 CCAAAGCATTTACTGGTATTATCAACCGTTTGGAGAAAGGACAAAGCGCGGATA 61

Qy 150 TyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleile--- 168
Db 62 TATGAAAGAAAGTGGCACCACTGTTACCTTAGCCTCCATCACCCTGACATCATCGGT 121

Qy 169 -----ValThrGluHisAlaAsnGlnAlaLysGlu 178
Db 122 AACAAATTTCTATGTTAACTATGTAGAAAGTAAATTAATCTGAACATGCTAATCAGGCCAAGGAG 181

Qy 179 ThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 198
Db 182 ACTCTGTATGAGATTACATAGACAAATACGACGGCATCGTCTGTGTGCGGGAGATGGT 241

Qy 199 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValasp 218
Db 242 ATGTTACGAGGAGTGTGCACGGTCTGATTTGGGAGGACGACAGGAGCGCGGGTCCAC 301

Qy 219 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAla 238
Db 302 CAGAACACCCCGGGCTGCTGGTCCCAAGTAGCTCCGATTCGGATTCATTTCCGCGCA 361

Qy 239 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 258
Db 362 GGGTCAACCGACTCGGTGTGTACTTCCACCGTGGGACCCAGCAGCAGAAACCTCGCG 421

Qy 259 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 278
Db 422 CTGCATATCGTTGTTGGGACTCGCTGGCCATGGATGTGCTCTCAGTCCACCAACACAGC 481

Qy 279 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleileys 298
Db 482 ACACCTCTCTCTACTCCGTGCTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAAG 541

Qy 299 AspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThr 318
Db 542 GACAGTGAGAAAGAACGGTGGTGGTCTTCCAGATACGACTTTTTCAGGTTTAAAGACC 601

Qy 319 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 338
Db 602 TTCTCTCTCCACCACTGCTATGAAAGGACAGTGTCTTCTCTCTCCCTGCACACACACAGGTG 661

Qy 339 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerIys 358
Db 662 GGATCTCCAAAGGATAGGAAGCCCTGCGGCGAGGATGCTTTGTTTTCAGGCAAGCAAG 721

Qy 359 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 378
Db 722 CAGCAGCTGGAGGAGGACAGAGAACATCTGTATGGTTTGGAGCTGCGGAGGAGCGTG 781

Qy 379 GluGluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaIleThrAsnMetSerCys 398
```

```
QY 239 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 258
DB 362 GGGTCAACGGACTGGTGTGTACTCCACCGTGGGCACGACGACGAGAACCTCGCGC 421
QY 259 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHisAsnSer 278
DB 422 CTGCATATCGTGTGGGACTCGTGGCCATGGATGTCTCTAGTCCACCAACACAGC 481
QY 279 ThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleLeuys 298
DB 482 ACACCTCTCGTACTCCGCTGCTGTGGCTACGGCTTCTACGGGGACATCATCAAG 541
QY 299 AspSerGluLysLysArgTyrPheLeuAlaArgTyrAspPheSerGlyLeuLysThr 318
DB 542 GACAGTGAAGAAGAACGGTGGTCTTGGCATACGACTTTTTCAGGTTTAAAGACC 601
QY 319 PheLeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrVal 338
DB 602 TTCCTCTCCACCATGCTATGAGGGACAGTGTCTTCTCTCCCTGCAACACACAGGTG 661
QY 339 GlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLys 358
DB 662 GGATCTCAAGGATAGGAAGCCCTGCCGGCAGGATGCTTTGTTGAGGCAAGCAAG 721
QY 359 GlnGlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal 378
DB 722 CAGCAGCTGGAGGAGGAGCAGAAAGCACTGTATGTTTGGAAAGCTGCCGAGGACGTG 781
QY 379 GluGluTyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCys 398
DB 782 GAGGAGTGGCAAGTCTGTGGGAAGTTTCTGGCCATCAATGCCAACCAATGTCTCT 841
QY 399 AlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSer 418
DB 842 GCTTGTCCCGAGAGCCCGAGGGCTCTCCCGGGCTGCCCACTTGGGAGACGGGTCTCT 901
QY 419 AspLeuIleLeuIleArgLysCysSerArgPheAsnPheLeuArgPheLeuIleArgHis 438
DB 902 GACCTCATCTCATCCGAAATGCTCCAGGTTCAATTTTCTGAGATTTCTCATCAGGCAC 961
QY 439 ThrAsnGlnGlnAspGln 444
DB 962 ACCAACGACGAGGACGAG 979
```

## RESULT 13

US-10-635-247

; Sequence 247, Application US/10115635

; Publication No. US20040137434A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Zhou, Ping

; APPLICANT: Goodrich, Ryle

; APPLICANT: Liu, Chenghua

; APPLICANT: Asundi, Vinod

; APPLICANT: Ren, Feiyan

; APPLICANT: Zhang, Jie

; APPLICANT: Zhao, Qing A.

; APPLICANT: Xue, Aidong J.

; APPLICANT: Yang, Yonghong

; APPLICANT: Weinman, Tom

; APPLICANT: Drmanac, Radoje T.

; TITLE OF INVENTION: Novel Nucleic Acids and

; FILE REFERENCE: Polypeptides

; FILE REFERENCE: 797CON

; CURRENT APPLICATION NUMBER: US/10/115,635

; CURRENT FILING DATE: 2002-04-03

; PRIOR APPLICATION NUMBER: 09/714,936

; PRIOR FILING DATE: 2000-11-17

; NUMBER OF SEQ ID NOS: 362

; SOFTWARE: pt\_FL\_genes Version 2.0

; SEQ ID NO 247

; LENGTH: 817

## ; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (90)..(662)

US-10-115-635-247

## Alignment Scores:

Pred. No.: 1,29e-111 Length: 817

Score: 1032.00 Matches: 190

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 35.73% Indels: 0

DB: 19 Gaps: 0

US-10-631-958-10 (1-537) x US-10-115-635-247 (1-817)

```
QY 1 MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuTyrValLysGlnGlnArg 20
DB 90 ATGGGGGGGACGGGGGGGGGGGAGCCGCTGCATCCGCTGCTGGTGGAGACGAGCGC 149
QY 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTyrTrpArgSerProGlyPro 40
DB 150 TGGCCGCTGAGCTGGAGCCCGCGCGGCTCTGCTGCGCTGCTGAGATCATCGCCGTT 209
QY 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal 60
DB 210 GGAGCCGCGCGCCCGCGCGGATGCCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 269
QY 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTyrGlnLysMetGluLys 80
DB 270 GAGGAACAGACGTTACGGGAAACATCAAGGCAGTGGAAATGGCAAAATGGAAAG 329
QY 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgArgHisArgTyrLysTrpAla 100
DB 330 CCTTACGCTTTTACAGTTCACTGTGTAAAGAGACGACGACGCGCGCTGGAAGTGGCG 389
QY 101 GlnValThrPheTyrCysProGluGlnLeuCysHisLeuTyrLeuGlnThrLeuArg 120
DB 390 CAGGTGATTTTCTCGTGTCCAGAGGACGAGTGTGTCACTTGTGTGCTGCAGACCTGCGG 449
QY 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
DB 450 GAGATGCTGGAGAAGCTGACGTCCAGACCAAGCATTTTACTGGTATTTTATCAACCCGTT 509
QY 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
DB 510 GGAGAAAAGGACAAAGCAAGCGGATATATGAAAGAAAAGTGGCACCCTGTTACCTTA 569
QY 161 AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180
DB 570 GCCTCCATCACCCTGACATCATCTTACTGAAACATGCTATACAGGCCAAGGAGACTCTG 629
QY 181 TyrGluIleAsnIleAspLysTyrAspGly 190
DB 630 TATGAGATTAAATAGACAAATATACGACGG 659
```

## RESULT 14

US-09-969-896-8

; Sequence 8, Application US/09969896

; Publication No. US2003012553A1

; GENERAL INFORMATION:

; APPLICANT: Kossida, Sophia

; TITLE OF INVENTION: Regulation of human Sphingosine

; FILE REFERENCE: Kinase-Like Protein

; FILE REFERENCE: 004974.00594

; CURRENT APPLICATION NUMBER: US/09/969,896

; CURRENT FILING DATE: 2001-10-04

; PRIOR APPLICATION NUMBER: US 60/238,005

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: US 60/314,113

; PRIOR FILING DATE: 2001-08-23

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FastSEO for Windows Version 4.0

REF ID	NO	8

LENGTH: 550

;  
:  
: TYPE: DNA

```

: LIFE: DNA
: ORGANISM: Homo sapiens

```

US-09-969-896-8

Alignment Scores:

Fragment Count:	8,498-93	Length:	550
Pred. No.:	871.00%	Matches:	158
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	30.16%	Gaps:	0
DB:	10		

US-10-631-958-10 (1-537) x US-09-969-896-8 (1-550)

Qy	1	MetGlyAlaThrGlyAlaAaGluProLeuGlnSerValLeuTrpValIysGlnGlnArg	20
Db	76	ATGGGGCGACGGGGCGGGCGGAGCCGCTGCAATCCGTGCTTGGGTGGAGCAGCAGCGC	135
Qy	21	CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro	40
Db	136	TGCGCGGTGAGCTGGAGCCCGCGGGGCTCTGCTGCGTGGTGGGGAGACCCCGGGGCC	195
Qy	41	GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleIleAlaVal	60
Db	196	GGAGCCGGCGCCCGCGCGGATGCTGCTCTGTGCTGTATCTGAGATCATCGCGGTT	255
Qy	61	GluGluThrAspValHisGlyIysHisGlnClySerGlyIysTrpGlnIysMetGluLys	80
Db	256	GAGGAAACAGACGTTTCAGCGGAAACATCAAGCGCAGTGGAAAAATGGCAAAATGGAAAAAG	315
Qy	81	ProTyrAlaPheThrValHisCysValIysArgAlaArgArgHisArgTrpIysTrpAla	100
Db	316	CCTTAGCGCTTTACAGTTTCATGTGTAAAGAGACACGACGGCACCGCTGGAAGTGGCG	375
Qy	101	GlnValThrPheTrpCysProGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg	120
Db	376	CAGGTGACTTCTTGCTGTCCAGAGGAGCAGCTGTGTCACTTGTGGCTGCAAGACCCCTGGCG	435
Qy	121	GluMetLeuGluLysLeuThrSerArgProIysHisLeuLeuValPheIleAsnProPhe	140
Db	436	GAGATGCTGGAGAACTGACGTCNAGCCAAAGCATTTACTGGTATTTATCAACCCGTTT	495
Qy	141	GlyGlyLysGlyGlnGlyLysArgIleTrpGluArgLysValAlaProLeuPhe	158
Db	496	GAGGAAAAAGCAACAGGCAAGCGGATATCAAAAAAAGTGGCAACACTGTTC	549

## RESULT 15

KEYS TO THE  
IIS-10-637-958-A

03-10-631-938-8  
: Sequence 8 Application US/10631958

; sequence 8, Application OS/1063  
: Publication No. IUS20040192580A1; PUBLICATION NO: US20  
: GENERAL INFORMATION:

; GENERAL INFORMATION:

; APPLICANT: KOBBIDA, SOPHIA  
: TITLE OF INVENTION: Regulation of human sphingosine

;	TITLE OF INVENTION: REGULATION OF HUMAN
:	KINASE-LIKE PROTEIN

; TITLE OF INVENTION: KINASE-  
; FILE REFERENCE: 004974 00594

FILE REFERENCE: 004374.00394  
CURRENT APPLICATION NUMBER: US/10/631 958

; CURRENT APPLICATION NUMBER: US/I  
 : CURRENT FILING DATE: 2003-08-01

; CURRENT FILING DATE: 2003-08-01  
 ; PRIOR APPLICATION NUMBER: US/09/969.896

; PRIOR FILING DATE: 2001-10-0

; PRIOR FILING DATE: 2001-10-04  
: PRIOR APPLICATION NUMBER: US 60/238,005

;  
;  
;

PRIORITY APPLICATION NUMBER: US 60/314,113

; PRIOR AFFIDAVIT NUMBER  
 : PRIOR FILING DATE: 2001-

: NUMBER OF SEO ID NOS: 16

: SOFTWARE: FastSEO for Windows Version 4.0

8 NO ID IP SEQ :  
8 NO ID IP SEQ :  
8 NO ID IP SEQ :

LENGTH: 550

```

: SERO1...: 3
: TYPE: DNA

```

ORGANISM: Homo sapiens

US-10-631-958-8

**Alignment Scores:**

Pred. No.:	8,498-93	Length:	550
Fragment Sources:		Matches:	158
Score:	871.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	30.16%	Gaps:	0
DB:	19		

US-10-631-958-10 (1-537) x US-10-631-958-8 (1-550)

Qy	20	1	MetGlyAlaThrGlyAlaAaGluProLeuGlnSerValLeuTrpValIleLeuAlaVal	Arg	20
Db	135	76	ATGGGGCGGACGCGGGCGCGGAGCGCGTGCATTCGTTGTGTGGGTGAACAGACAGCGC		135
Qy	40	21	CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro		40
Db	195	136	TGGCGCGTGAGCCTGGAGCCCGCGGGCTCTGTGGCTGGTGGCGAGCCCGGGGCC		195
Qy	60	41	GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleIleAlaVal		60
Db	255	196	GGAGCGGCGGCCCGCGCGCGATGCTGTCTGTGCTGTATCTGAGATCATCGCGTT		255
Qy	80	61	GluGluThrAspValHisGlyIleHisGlnGlySerGlyIleTrpGlnIleMetGluIle		80
Db	315	256	GAGGAACACAGACGTTTCAGCGGAAACATCAAGCGCAGTGGAAAATGGCAGAAAATGGAAAAG		315
Qy	100	81	ProTyrAlaPheThrValHisCysValIysArgAlaArgHisArgTrpIleTrpAla		100
Db	375	316	CCTTACGCTTTTACAGTTTCATGTGTAAAGAGAGACGACGGCACCGCTGGAGTGGCG		375
Qy	120	101	GlnValThrPheTrpCysProGluGlnLeuCysHisIleuTrpLeuGlnThrLeuArg		120
Db	435	376	CAGGTGACTTTCTGTGTCTCAGAGAGCAGCTGTGTCTACTTGTGGCTGCGAGACCTGCGG		435
Qy	140	121	GluMetLeuGluIleLeuThrSerArgProIleHisIleLeuValPheIleAsnProPhe		140
Db	495	436	GAGATGCTGGAGAACTGACGTCCAGACCAAGCAATTTATCTGTATTATTCAACCCGTTT		495
Qy	158	141	GlyGlyIleGlyGlnGlyIleArgIleTyrGluArgIysValAlaProLeuPhe		158
Db	549	496	GAGGAAAGGACCAAGCAGCGGATATGAAAGAAAGTGGCACCACCTGTTTC		549

Search completed: September 6, 2005, 10:48:14

Job time : 1559.64 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 5, 2005, 19:28:44 ; Search time 303.735 Seconds  
(without alignments)  
2892.921 Million cell updates/sec

Title: US-10-631-958-10  
Perfect score: 2888  
Sequence: 1 MGATGAEPQLSVLWVKQR.....QLVRLFARGIENKPDVSHS 537

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues  
Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US10631958/runat\_02092005\_165811\_4432/app\_query.fasta\_1.1941  
-DB=Issued Patents NA -OPMT=fastap -SUPFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -SPART=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10631958 @CGN\_1\_1\_397 @runat\_02092005\_165811\_4432 -NCFU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/FACTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	ID Description
1	2881	99.8	4432 4 US-09-774-528-148 Sequence 148, App
2	801.5	27.8	2064 4 US-09-270-767-14306 Sequence 14306, A
3	585.5	20.3	1084 4 US-09-270-767-15155 Sequence 15155, A
4	302.5	10.5	901 4 US-09-270-767-30448 Sequence 30448, A
5	278	9.6	1857 4 US-09-970-516-3 Sequence 3, Appli
6	278	9.6	2380 4 US-09-817-676A-13 Sequence 13, Appl
7	269	9.3	2598 4 US-09-817-676A-11 Sequence 11, Appl
8	266.5	9.2	1205 4 US-09-959-897-1 Sequence 1, Appli
9	264.5	9.2	1155 4 US-09-970-516-1 Sequence 1, Appli
10	263.5	9.1	1783 4 US-09-949-016-1155 Sequence 1155, Ap
11	257.5	8.9	1533 4 US-09-205-258-90 Sequence 90, Appl
12	247.5	8.6	1149 4 US-09-970-516-5 Sequence 5, Appli

13	228	7.9	1611	4	US-09-248-796A-1756	Sequence 1756, Ap
14	220.5	7.3	1875	4	US-09-614-221A-399	Sequence 399, App
15	203.5	7.0	1764	4	US-09-949-016-3940	Sequence 3940, App
16	199.5	6.9	2462	4	US-09-620-312D-796	Sequence 796, App
17	184	6.4	1012	4	US-09-270-767-12677	Sequence 12677, A
18	153	5.3	942	4	US-09-107-532A-3500	Sequence 3500, Ap
19	153	5.3	1239	4	US-09-603-208A-263	Sequence 263, App
20	144	5.0	959	4	US-09-270-767-6490	Sequence 6490, Ap
21	144	5.0	959	4	US-09-270-767-21772	Sequence 21772, A
22	127	4.4	7195	4	US-09-949-016-12897	Sequence 12897, A
23	127	4.4	7198	4	US-09-949-016-15682	Sequence 15682, A
24	125	4.3	485	4	US-09-270-767-31476	Sequence 31476, A
25	123.5	4.3	915	4	US-09-134-000C-1961	Sequence 1961, Ap
26	119.5	4.1	3089	4	US-09-634-238-188	Sequence 188, App
27	118	4.1	912	4	US-09-107-532A-1479	Sequence 1479, Ap
28	118	4.1	4411529	3	US-09-103-840A-1	Sequence 1, Appli
29	116.5	4.0	885	4	US-09-583-110-1079	Sequence 1079, Ap
30	116.5	4.0	948	4	US-09-107-433-518	Sequence 518, App
31	113.5	3.9	1035	4	US-09-134-000C-408	Sequence 408, App
32	113	3.9	2925	4	US-09-883-134-6	Sequence 6, Appli
33	112.5	3.9	2612	4	US-09-949-016-4164	Sequence 4164, Ap
34	111.5	3.9	900	3	US-08-961-527-332	Sequence 332, App
35	111.5	3.9	93398	4	US-09-949-016-14167	Sequence 14167, A
36	110	3.8	975	3	US-09-134-001C-1937	Sequence 1937, Ap
37	110	3.8	3761	4	US-09-710-279-3656	Sequence 3656, Ap
38	110	3.8	10717	4	US-09-902-540-991	Sequence 991, App
39	109.5	3.8	36941	4	US-08-311-731A-130	Sequence 130, App
40	109	3.8	498	4	US-09-893-737-59	Sequence 59, Appl
41	107.5	3.7	978	4	US-09-107-532A-3132	Sequence 3132, Ap
42	107	3.7	191	4	US-09-270-767-28485	Sequence 28485, A
43	106	3.7	4411529	3	US-09-103-840A-1	Sequence 1, Appli
44	105.5	3.7	2469	4	US-09-252-991A-5908	Sequence 5908, Ap
45	105.5	3.7	3135	4	US-09-252-991A-5922	Sequence 5922, Ap

ALIGNMENTS

RESULT 1  
US-09-774-528-148  
; Sequence 148, Application US/09774528  
; Patent No. 6743619  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Zhou, Ping  
; APPLICANT: Goodrich, Ryle  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Zhang, Jie  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6743619el Nucleic Acids and  
; FILE REFERENCE: 802  
; POLYPEPTIDES  
; CURRENT APPLICATION NUMBER: US/09/774, 528  
; CURRENT FILING DATE: 2001-01-30  
; NUMBER OF SEQ ID NOS: 441  
; SOFTWARE: pt\_FL\_genes Version 2.0  
; SEQ ID NO 148  
; LENGTH: 4432  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1129) ..(2817)  
US-09-774-528-148  
Alignment Scores:

Pred. No.:	1.19e-317	Length:	4432
Score:	2881.00	Matches:	536
Percent Similarity:	99.81%	Conservative:	0
Best Local Similarity:	99.81%	Mismatches:	1
Query Match:	99.76%	Indels:	0
DB:	4	Gaps:	0
US-10-631-958-10 (1-537) x US-09-774-528-148 (1-4432)			
Qy	1	MetGlyAlaThrGlyAlaAlaGluProLeuGlnSerValLeuThrPValLysGlnGlnArg	20
Db	1204	ATGGGGCGACGGGGCGGGAGCCGCTGCAATCCGTCGTGTGGGTGAACAGCAGCGC	1263
Qy	21	CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro	40
Db	1264	TGCGCCGTGAGCTTGAGACCGCGCGGGCTCTGCTGCGCTGGTCGGGAGCCCGGGGCC	1323
Qy	41	GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluLleIleAlaVal	60
Db	1324	GGAGCCGGGCCCCCGGGCGGATGCTCTGTGCTGTATCTGAGATCATCGCCGTT	1383
Qy	61	GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys	80
Db	1384	GAGGAACAGACGTTTACGGGAACATCAGGCCAGTCGAAATCGCAGAAATCGAANAAG	1443
Qy	81	ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAla	100
Db	1444	CCTTACGCTTTTACAGTTTCACTGTGTAAAGACACGACGCCACCGCTGGAAGTGGCGC	1503
Qy	101	GlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg	120
Db	1504	CAGGTGACTTTCTGGTGTCCAGAGAGCAGCTGTCTACTTGTGGCTGACAGACCCTCGGG	1563
Qy	121	GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe	140
Db	1564	GAGATGCTGGAGAAAGCTGAGCTCCAGACCAAGCATTTACTGTGTATTTATCAACCCGTTT	1623
Qy	141	GlyGlyLysGlnGlnLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu	160
Db	1624	GGAGGAAAGGACAAAGCAAGCGGATATATGAAAGAAAGTGGCACCACCTGTTTCCACCTTA	1683
Qy	161	AlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu	180
Db	1684	GCCTTCATACCATGACATCATCTTATGAACATGCTAATCAGGCCAAGGAGACTCTG	1743
Qy	181	TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe	200
Db	1744	TATCAGATTAAACATAGACAAATACGCGGCATCGTCTGTGTCCGGGAGATGTTATGTTTC	1803
Qy	201	SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn	220
Db	1804	AGCCAGGTGCTGCACGTCTGTATTGGAGAGACGACAGAGAGCGCGGGGTGACACAGAAC	1863
Qy	221	HisProArgAlaValLeuValProSerSerLeuArgIleGlyLleIleProAlaGlySer	240
Db	1864	CACCCCGGGCTGTGCTGGTCCCAGTAGCCTCCGGATTGGAATCATCTCCGACGGGTCA	1923
Qy	241	ThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis	260
Db	1924	ACGGACTGCGTGTGTTACTCCACCGTGGGCACACGACGAGCAGAAAACTCGCGCCTGCAT	1983
Qy	261	IleValValGlyAspSerLeuAlaMetAspValSerSerValHisAsnSerThrLeu	280
Db	1984	ATCGTGTGGGGACTCGCTGGCCATGGATGTGTCCTCAGTCCACCAACAACAGACACTC	2043
Qy	281	LeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSer	300
Db	2044	CTTCGCTACTCCGTGTCCCTGCTGGGCTACGGCTTCTACGGGGACATCATCAGGACAGT	2103
Qy	301	GluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeu	320
Db	2104	GAGAAGAAACCGTGGTGGGTCTTGGCAGATACACACTTTTTCAGGTGTAAAGACCTTCCTC	2163

Qy	321	SerHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySer	340
Db	2164	TCCACCACTGCCTATGGAGGGACAGTGTCTCTCTCCCTGCACAAACACACGGTGGGATCT	2223
Qy	341	ProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGln	360
Db	2224	CCAAAGGATAGGAAGCCCTCCGGCGAGGATGCTTTGTTTGACGGCAAAACNAGCAGCAG	2283
Qy	361	LeuGluGluGluGlnIlystLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGluGlu	380
Db	2284	CTGGAGGAGCAGCAGAAAGAACTGTATGTGTTTGGAAAGCTGCGAGGACGTGGAGGAG	2343
Qy	381	TyrGlnValValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCys	400
Db	2344	TGGCAAGTCGTCTGTGGGAAGTTTCTGGCCATCAATGCCCAAAACATGTCTCTGTGCTTGT	2403
Qy	401	ArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeu	420
Db	2404	CGCCGAGCCCCAGGGGCGCTCTCCCGGCTGCCCACTTGGGAGACGGGTCTTCTGACCTC	2463
Qy	421	IleLeulleArgIlyCysSerArgPheAsnPheLeuArgPheIleulleArgHisThrAsn	440
Db	2464	ATCCTCATCCGAAATGCTCCAGGTTCAAATTTCTTGAGATTTCTCATCAGGCAACACCAAC	2523
Qy	441	GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe	460
Db	2524	CAGCAGACAGATTGACTTCTTTGTTTGAAGTTTATCGCTCAAGAAATTCAGTTT	2583
Qy	461	ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyLysLysArgPhe	480
Db	2584	ACGTGGAAGCACATGGAGGATGAGGACAGCCACTCAAGGAGGGGGGAAAGACGCTTT	2643
Qy	481	GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTroAsn	500
Db	2644	GGGCACATTTCAGCAGCCACCCCTCTGCTGTGCACCGTCTCCAAACAGCTCCTCGAAC	2703
Qy	501	CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal	520
Db	2704	TGGCAGGGGAGGTCTCTGCACAGCCCTGCCATCGAGGTCAGAGTCCACTCCAGCTGTT	2763
Qy	521	ArgLeuPheAlaArgIlyIleGluGluAsnProLysProAspSerHisSer	537
Db	2764	CGACTCTTTCAGGAGGAATTGAAGAAATCCGAAGCCAGACTCACACAGC	2814
RESULT 2			
US-09-270-767-14306			
; Sequence 14306, Application US/09270767			
; Patent No. 6703491			
; GENERAL INFORMATION:			
; APPLICANT: Homburger et al.			
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster			
; FILE REFERENCE: File Reference: 7326-094			
; CURRENT APPLICATION NUMBER: US/09/270,767			
; CURRENT FILING DATE: 1999-03-17			
; NUMBER OF SEQ ID NOS: 62517			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 14306			
; LENGTH: 2064			
; TYPE: DNA			
; ORGANISM: Drosophila melanogaster			
US-09-270-767-14306			
Alignment Scores:			
Pred. No.:		1.02e-80	Length: 2064
Score:		801.50	Matches: 197
Percent Similarity:		46.74%	Conservative: 97
Best Local Similarity:		31.32%	Mismatches: 206
Query Match:		27.75%	Indels: 129
DB:		4	Gaps: 16
US-10-631-958-10 (1-537) x US-09-270-767-14306 (1-2064)			

11 GlnSerValLeuTrpValLysGlnGlnArqCysAlaValSerLeuGlu-----ProAla 28

```
Db 175 CAGCAGTTGGTTGGAGCGTCTGCAGAAAGATCAAGCAATGCCCAAGGAACAGAGGCC 234
Qy 29 ArgAlaLeuLeuArgTrpTrpArgSerProGlyProGlyAlaGlyAlaProGlyAlaAsp 48
Db 235 AAGGACACCGTTG-----CGCGCGGACTCGCCAGCTCAACACGAGGCGGAATC 279
Qy 49 AlaCysSerValProValSerGluIleLeuAlaValGluGluThrAspValHisGlyLys 68
Db 280 TGCTCTTATGGACCAAGACCGCATCTGCTCCACCTGGAC-----GATGTGTTAGCATA 333
Qy 69 HisGlnGlySerGlyLysTrpGlnLysMetGluLysPro----- 81
Db 334 AGGACGGGGACACCAAGCGAGCTCCCTCAAAACCGCTTCGCCCGCAGGCTCGGAACGC 393
Qy 82 -----TyrAlaPheThrVal 86
Db 394 AGTTCGGGATGCAGTGGCGATGTTGCTCAAGGGAAGCCTACAGTCAATACCTGACCATC 453
Qy 87 HisCysValLysArgAlaArg-----HisArgTrpLysTrpAlaGln 101
Db 454 AACTATGCCATGAGGCTGAGCAAGTCCCAACGGAATTCGATCGGTGGAGCTAAGGAGA 513
Qy 102 ValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArgGlu 121
Db 514 CTCACCTTCTTCAACTCGGATCCCTACATAGTAGGCGAGTGGGACCAAGGAGCTGCAATC 573
Qy 122 MetLeuGluLysLeuThr-----SerArgProLysHisLeuLeuValPheIleAsn 138
Db 574 CGTCTGCATAGCTCTCCCGCCACCGGATGCGAGTGGCGCGCTTTTGCTTTATATAAC 633
Qy 139 PropheGlyGlyLeuGlyGlnGlyLysArgIleTrpGluArgLysValAlaProLeuPhe 158
Db 634 CCCTATGAGGTGCGAAGCGGAGCTCAGACCTATGAGCGCATGTGAGACCTATTATTC 693
Qy 159 ThrLeuAlaSerIleThrAspIleLeuValThrGluHisLeuAlaAsnGlnAlaLysGlu 178
Db 694 CAGCTTGCGGCTAGAGCCACGCTGTATCACCTCAGAGGGCAACCAAGTGAAGGAC 753
Qy 179 ThrLeuTyrgluIleAsnLysLysTyrsArgGlyIleValCysValGlyLysAspGly 198
Db 754 ATACTTCTGAGCCATGATCGGAGTATACGATGCGGTTGCTGTGTCGGAGCGATGCG 813
Qy 199 MetPheSerGluValLeuHisGlyLeuLeuGlyArgThrGlnArgSerAlaGlyValAsp 218
Db 814 ACCGTAGCAGAGGTCAATCAACGCACTGATATTCGCTCAAAATGCGAGAGTTGGGACTGAC 873
Qy 219 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleProAla 238
Db 874 GAACAGCGGCCACCATACATTCCTCAAGACCG---GCTCTGCCAGTGGGTGATTCGCGCT 930
Qy 239 GlySerThrAspCysValCysTyrsSerThrValGlyThrSerAspAlaGluThrSerAla 258
Db 931 GGCAGACCGACACCATTCGCTATAGTATGACGCGCAGCGGCGGATGTGAGNACAGCGCT 990
Qy 259 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisAsnSer 278
Db 991 ATCCATGTGATTCGTGGCGCAGCATCGGGATTTGGATGTGTGCAGTGTGAGCAATGGCGAC 1050
Qy 279 ThrLeuLeuArgTyrsSerValSerLeuLeuGlyTyrglyPheTyrglyAspIleIleLys 298
Db 1051 TCCCTGCTCAGATTCTGTGCCAGTGTCTGAGCTACGGGTACCTCGGGCGATGTGGCAGCC 1110
Qy 299 AspSerGlyLysArgTrpLeuGlyLeuAlaArgTyrsAspPheSerGlyLeuLysThr 318
Db 1111 CAGAGCGAGAATACCTCGTGGATGGAGCCGCCCGGTACGAGTACAGTGGCGTCAAGGCC 1170
Qy 319 PheLeuSerHisGlyCysTyrgluGlyThrValSerPheLeu----- 332
Db 1171 TTCCTGAATATCGCGGCTATCAGCCGCAACTGAGAAATGTTAGAAGAGCCCGCATCTTCTA 1230
Qy 333 -----ProAlaGlnHisThrValGlySerProArgAsp----- 343
```

```
Db 1231 CTGACCAGCGCGCTGGAGGACATTTCCGACAGTCCGGATAGTGTGCTCGCTGGGAGAG 1290
Qy 344 -----ArgLysProCysArgAlaGlyCysPheValCys----- 354
Db 1291 TCAGTGGCATCCGCTCTGATGCCAATTTGCCAGCGCTGCAGCTTCGCCAGCAGCATACAG 1350
Qy 355 ---ArgGlnSerLysGlnGlnLeuGluGluGlnLysValAlaLeuTyrglyLeuGlu 373
Db 1351 GAACAGCGATCCTCAITTTGTCATCCAAAGAGAACTCTAAAGAGGCA-----GAG 1398
Qy 374 AlaAlaGluAspValGlu----- 379
Db 1399 CGCAATCAGCAGGTAGAAAACAGAGGACTCTCATCTAGCGCCGACGAGCAGCAGCTCTG 1458
Qy 379 ----- 379
Db 1459 AGGCCTCGTCCGCTCCAGGCAATTTTCGATTGGCCACTGGCTCCATTTTCATCAATGAGG 1518
Qy 380 -----GluTrpGlnValValCysGlyLysPheLeuAlaIleAsnAlaThr 394
Db 1519 AACCTCGGCAACGATCAGTGGAGGTTGTGGGGGCAATTTCTTTATGATCTGCGGCGCG 1578
Qy 395 AsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGly 414
Db 1579 AACATACTCGGCTCGGCCAGGAGTCCCAATGGCATCTCCGTTACAGTCATCTGGGT 1638
Qy 415 AspGlySerSerAspLeuIleLeuLeuArgLysCysSerArgPheAsnPheLeuArgPhe 434
Db 1639 GATGTTGCTCGTGGACCTGATTTCTCGTGAAGAACCTCCTCCTCAACAACGCTGGGTTT 1698
Qy 435 LeuIleArgHisThrAsnGlnGln---AspGlnPheAspPheThrPheValGluValTy 453
Db 1699 CTGCTCAACAGCGCGGCGAGAAAGTGGTGATATTCGCAATTTGCTTTGTAGAGGTATAT 1758
Qy 454 ArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeuLys 473
Db 1759 CGCACAGGGAATTCGGTTCAGAACTTTTCTCCAGCGAGGAGGACTACAGCTTGCA 1818
Qy 474 GluGlyGlyLysLysArgPheGlyHisIleCys----- 484
Db 1819 GGGTCT-----TGTCAGCGCGATTACACCGCTGAAGAAATG 1854
Qy 485 SerSerHisProSerCysCysThrValSerAsnSerSerTrpAsnCysAspGlyGlu 504
Db 1855 ACCGCCCACTCTCC-----TCGACGAGGATTTTCCAGCTGGAACCTGTGATGGAGAG 1905
Qy 505 ValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuValArgLeuPheAla 524
Db 1906 GTGTGACCGNACCTGGACATACCATGGGATCGCATTTGTCACTCATCGAGGTCTTCATG 1965
Qy 525 ArgGlyIleGluGluAsnProLysPro 533
Db 1966 CGGGGTCCCCATCTCTACAGCAAGCCC 1992

RESULT 3
US-09-270-767-15155/c
; Sequence 15155, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270.767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15155
; LENGTH: 1084
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-15155

Alignment Scores:
```



```
Pred. No.: 1,56e-56 Length: 1084
Score: 585.50 Matches: 125
Percent Similarity: 51.14% Conservative: 55
Best Local Similarity: 35.51% Mismatches: 103
Query Match: 20.27% Indels: 69
DB: 4 Gaps: 6

US-10-631-958-10 (1-537) x US-09-270-767-15155 (1-1084)

Qy 129 ArgProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyArg 148
Dy 1044 CGAGTGGCCGCCCTTTTGTCTTTATATAAACCCCTATGGAGGTTCGAAAGCGGGAGCTCAG 985
Qy 149 IleTyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIle 168
Dy 984 ACCTATGAGCCCATGTGAGACCTATTTCAGCTTCCGCGGTAGAGCCACCGTGTATC 925
Qy 169 ValThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyr 188
Dy 924 ACCACTCAGAGGGCAAAACCAAGTGAAGACATACCTTCTGAGCCATGATCTGGGAGTATAC 865
Qy 189 AspGlyIleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIle 208
Dy 864 GATCGGTTTCTGTCTCGAGGGCATGGCACCGTACGAGAGGTATCAACGGAAGTATGATA 805
Qy 209 GlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValPro 228
Dy 804 TTCGCTCAATGCGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 745
Qy 229 SerSerLeuArgIleGlyIleProAlaGlySerThrAspCysValCysTyrSerThr 248
Dy 744 --GCTTCGCGAGTGGTGTATTCCTCCGCTGGCAGCACCGACACCATTCGCTATAGTATG 688
Qy 249 ValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAla 268
Dy 687 CACGCGACGGCGATGTGAGGACACGCGCTATCCATGTGATTCTGGGCGACATCGGGGA 628
Qy 269 MetAspValSerValHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeu 288
Dy 627 TTGATGTGTGAGTGTGAGCAATGGCCAGTCCCTGCTCAGATCTGTGCGCAGTGTCTGT 568
Qy 289 GlyTyrGlyPheTyrGlyAspIleLeuLysAspSerGluLysLysArgTyrLeuGlyLeu 308
Dy 567 AGCTACGGGTACCTGGCGGATGTGGCAGCCAGAGCGAGAACTACCGCTGATGAGCGCGAA 508
Qy 309 AlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThr 328
Dy 507 CGCCGGTACGAGTACAGTGGCGTCAAGGCTTCTCTGAATAATACCGGCTATGAGCGCGAA 448
Qy 329 ValSerPheLeu-----ProAlaGlnHisThrValGly 339
Dy 447 CTGAGATGTTAGAGAGCCGATCTTCTACTGACCACCGCTGGAGGACATTCGCGAG 388
Qy 340 SerProArgAsp-----ArgLysProCysArgAlaGlyCys 351
Dy 387 AGTCGGATAGTGTGTGCTGCTGGGAGAGTCAGTGCCATCCGCTGTCTATGCCAATTGC 328
Qy 352 PheValCys-----ArgGlnSerLysGlnGlnLeuGlu 363
Dy 327 CAGCGCTCAGCTTCGCCAGCAGCATACAGGAACAGCGATCTCATTTGTTATCCAGAG 268
Qy 364 GluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspValGlu----- 379
Dy 267 GAATCTAAGAGGCA-----GAGCGCAATCAGCAGGTAGAAACAGAGGACTCT 220
Qy 379 ----- 379
Dy 219 CATCTAGCGCGCAGTGAAGCAGCACTCTTAGGCTCTGTCGCGCTCCAGGCAATCTTCGA 160
Qy 380 -----GluTyrGlnValVal 384
Dy 159 TTGCCCACTGGCTCCATTTTCATCAATAGGAACCTTCGGCAACGATCAGTGAAGGTTGTG 100

Alignment Scores:
Pred. No.: 2,34e-24 Length: 901
Score: 302.50 Matches: 85
Percent Similarity: 41.24% Conservative: 35
Best Local Similarity: 29.21% Mismatches: 82
Query Match: 10.47% Indels: 89
DB: 4 Gaps: 9

US-10-631-958-10 (1-537) x US-09-270-767-30448 (1-901)

Qy 317 LysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPheLeu----- 332
Dy 2 AAGSCCTTCTGTAATATCGCGCTATGACCGCGAATGTAGAAAGAGCCCGAT 61
Qy 333 -----ProAlaGlnHisThrValGlySerProArgAsp----- 343
Dy 62 CTTCTACTGACACCGCGCTGAGGACATTCGCGAGATCCGATAGTGTGTCTCGTG 121
Qy 344 -----ArgLysProCysArgAlaGlyCysPheValCys----- 354
Dy 122 GGAGAGTCAGTGCATCCGCTGCTATGCCAATTCGCGAGCTGCAGCTTCGCCAGCAGC 181
Qy 355 -----ArgGlnSerLysGlnGlnLeuGluGlnLysLysAlaLeuTyrGly 371
Dy 182 ATACAGGAACAGCGATCTCTATTGTTTCATCAAGAGGAATCTAAAGAGCA----- 232
Qy 372 LeuGluAlaAlaGluAspValGlu----- 379
Dy 233 ---GAGCGCAATCAGCAGGTAGAAACAGAGACTCTCATCTAGCGCGCAGTGAAGCAGCA 289
Qy 379 ----- 379
Dy 290 CTCCTGAGGCTCTCGCGCTCCAGGCAATCTTCGATTGCGCCACTGGCTCCATTCATCA 349
Qy 380 -----GluTyrGlnValValCysGlyLysPheLeuAlaIleAsn 392
Dy 350 ATGAGGAACCTCGCGAACGATCAGTGAAGGTTGTGGGGCAATTTCTTTATGATCTGC 409
Qy 393 AlaThrAsnMetSerCysAlaCysArgSerProArgGlyLeuSerProAlaAlaHis 412
Dy 410 GGCGCGAACAATAACCTGCGCTGCGCAGGAGTCCCAATGGCATCTCCGTTACAGTCAT 469
Qy 413 LeuGlyAspGlySerSerAspLeuIleLeuIleArgLysCysSerAspPheAsnPheLeu 432
Dy 470 CTGGGTATGGTTCCTGGAGCTGATTTCTCGTGAAGAGACCTCAGTCTCAACACGCTG 529
Qy 433 ArgPheLeuIleArgHisThrAsnGlnGln----AspGlnPheAspPheThrPheValGlu 451
```

Db 530 CGTTTCTCTCAACACGGGGGAGAGTGGTATTCGCAATTTCCTTTTAGAG 589  
QY 452 ValTyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluSerAsp 471  
Db 590 GTATATCGCACAGAGGAATTCGGTTTCAGAACTTTTTCGCCAGCGAGGAGCTACAGC 649  
QY 472 LeuLysGluGlyGlyLysLysArgPheGlyHisIleCys----- 484  
Db 650 TTGGCAGGGTCT-----TGTCACCGCATTAACACGGCTGAA 685  
QY 485 -----SerSerHisProSerCysCysThrValSerAsnSerSerTrpAsnCysAsp 502  
Db 686 GAAATGACCGCCCACTCTCC-----TCGACGGAGTTTTCAGCTGGAACTGTGAT 736  
QY 503 GlyGluValLeuHisSerProLalleGluValArgValHisCysGlnLeuValArgLeu 522  
Db 737 GGAGAGGTGCTGACCGACCTGGACATAACCATCGCATTCATGTCAGCTCATCGAGGTC 796  
QY 523 PheAlaArgGlyIleGluGluAsnProLysPro 533  
Db 797 TTATGTCGGGGTCTCCCATTCCTACAGCAAGCCC 829

## RESULT 5

US-09-970-516-3  
; Sequence 3, Application US/09970516  
; Patent No. 6610534  
; GENERAL INFORMATION:  
; APPLICANT: NO. 6610534artis AG  
; TITLE OF INVENTION: Induction of blood vessel formation through administration of  
; FILE REFERENCE: 4-31617  
; CURRENT APPLICATION NUMBER: US/09/970,516  
; CURRENT FILING DATE: 2001-10-04  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 1857  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1857)  
; OTHER INFORMATION:  
US-09-970-516-3

Alignment Scores:  
Pred. No.: 5,08e-21 Length: 1857  
Score: 278.00 Matches: 91  
Percent Similarity: 44.24% Conservative: 55  
Best Local Similarity: 27.58% Mismatches: 155  
Query Match: 9.63% Indels: 30  
DB: 4 Gaps: 6

US-10-631-958-10 (1-537) x US-09-970-516-3 (1-1857)

QY 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTTrpTrpArgSerProGlyPro 40  
Db 128 TGGCGCCCAACCTGAAGCCAGGCCCGCGG---GTGGCTGTGCTCCGTGGCCGAGGTCT 184  
QY 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleIleAlaVal 60  
Db 185 CAGGCTGTCTGCACCC---TGCAGAGCCGCGAGCCCTCAGACTCAGCGGCCTACTTCTGCA 241  
QY 61 GluGluThrAspValHisGlyLysHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80  
Db 242 TCTAC-ACCTACCTCGGGCGCGCGGGCGCGAGAGCCACTCGCACCTTCGG 300  
QY 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgHisArgTrpLysTrpAla 100  
Db 301 CGAGATGGGGCGCCCACTACGAGAGAACCGTGGCGAGGCCCGAGCGCTGGGCCACTGCC 360  
QY 101 GlnValThrPheTrpCysProGluGluGlnLeuCysHisLeuTrpLeuGlnThrLeuArg 120

Db 361 CTCACCTGTCTG-----CTCCGAGACTGCCACTGCCCGGGGATGGG 402  
QY 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140  
Db 403 GAGATCACCCCTGACCTGCTACCTCGCGCGCCCGTGTCTTCTATTGGTCAATCCCTTT 462  
QY 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160  
Db 463 GGGGGTGGGGCTGGCGCTGGCAGTGGTGTAAAGAACCAACCATCTCCCATGATCTCTGAA 522  
QY 161 AlaSerIleThrThrAspIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180  
Db 523 GCTGGCTGTCTTCAACTCATCCAGACAGACAGAACCAACCGCCGGAGGTGGTTC 582  
QY 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200  
Db 593 CAGGGCTGAGCGCTGAGTGAGTGGATGCGTACGGTCTCGGAGAGCGGGTGTCTC 642  
QY 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220  
Db 643 CATGAGGTGCTGAACGGCTCTAGATGCG-----GTGGGATCTCCCTCGCGGTG 678  
QY 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240  
Db 679 TGGGAGGAAGCTGTGAAGATGCT-----GTGGGATCTCCCTCGCGGTG 726  
QY 241 ThrAspCysValCysTyrSer-----ThrValGlyThr 251  
Db 727 GGCAACGGCTGGCGGAGCAGTGAACAGCAGCGGGGATTTGACCCAGCCCTGGGGCTC 786  
QY 252 SerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspVal 271  
Db 787 GACCTGTGTCTCAACTGCTACTGTGCTGCGGGGTGGTGGCCACCCACCTGACCTG 846  
QY 272 SerSerValHisHisAsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGly 291  
Db 847 CTCTCCGTGACGCTGGCGCTCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 906  
QY 292 PheTyrGlyAspIleIleLysAspSerGluLysLysArgTrpLeuGlyLeuAlaArgTyr 311  
Db 907 TTCGTGTCTGATGTGATATCCAGAGCGAGCGCTTCAGGGCTTGGGAGTGGCCGCTTC 966  
QY 312 AspPheSerGlyLeuLysThrPheLeuSerHisCysTyrGluGlyThrValSerPhe 331  
Db 967 ACACCTGGGACGGTGTGGGCTCGCCACACTGACACACTACCGCGGAGCGCTCTCTCTAC 1026  
QY 332 LeuProAlaGlnHisThrValGlySerPro 341  
Db 1027 CTCCCGCCCACTGTGGAACCTGCTCGGCC 1056

## RESULT 6

US-09-817-676A-13  
; Sequence 13, Application US/09817676A  
; Patent No. 6800470  
; GENERAL INFORMATION:  
; APPLICANT: Spiegel, Sarah  
; APPLICANT: Kohama, Takafumi  
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,  
; FILE REFERENCE: 00170/HG  
; CURRENT APPLICATION NUMBER: US/09/817,676A  
; CURRENT FILING DATE: 2001-03-26  
; PRIOR APPLICATION NUMBER: US 60/194,318  
; PRIOR FILING DATE: 2000-04-03  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 2380  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS

```
; LOCATION: (7)..(1960)
; PUBLIC INFORMATION:
; TITLE: Molecular cloning and functional characterization of a
; TITLE: novel mammalian sphingosine kinase type 2 isoform
; JOURNAL: J. Biol. Chem.
; VOLUME: 275
; ISSUE: 26
; PAGES: 19513-19520
; DATE: 2000-06-30
; DATABASE ACCESSION NUMBER: AF245447
; DATABASE ENTRY DATE: 2000-06-27
US-09-817-676A-13

Alignment Scores:
Pred. No.: 7,78e-21 Length: 2380
Score: 278.00 Matches: 91
Percent Similarity: 44.24% Conservative: 55
Best Local Similarity: 27.58% Mismatches: 155
Query Match: 9.63% Indels: 30
DB: 4 Gaps: 6

US-10-631-958-10 (1-537) x US-09-817-676A-13 (1-2380)

Qy 21 CysAlaValSerLeuGluProAlaArgAlaLeuLeuArgTrpTrpArgSerProGlyPro 40
Db 134 TCGCGCCCAACCTGAAGCCAGCCGCCCGG---GTGGCTCGTCCGTTGGCGGAGGTCT 190
Qy 41 GlyAlaGlyAlaProGlyAlaAspAlaCysSerValProValSerGluIleAlaVal 60
Db 191 CAGGCTGTGCACCC---TGGAGAGCCGACCCCTCAGATCAGCGGCCCTACTTCTGCA 247
Qy 61 GluGluThrAspValHisGlyHisGlnGlySerGlyLysTrpGlnLysMetGluLys 80
Db 248 TCTAC-ACCTACCTCGGGCGCGCGGGCGCGCCGCGCCAGCACCTCGCACCTTCGG 306
Qy 81 ProTyrAlaPheThrValHisCysValLysArgAlaArgAGHisArgTrpLysTrpAla 100
Db 307 GCAGATGGGGCGCCACCTACGAGAGAACCGTGCAGGAGCCAGCGCTGGGCCACCTGCC 366
Qy 101 GlnValThrPheTrpCysProGluGlnGlnCysHisLeuTrpLeuGlnThrLeuArg 120
Db 367 CTCACCTGCTCG-----CTCGAGGAGCTGCCACTCCCGGGGATGGG 408
Qy 121 GluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsnProPhe 140
Db 409 GAGATCACCCCTGACCTGCTACCTCGGCGCGCCCGTTGCTTCTATTGGTCAATCCCTTT 468
Qy 141 GlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPheThrLeu 160
Db 469 GGGGTCGGGGCTCGCCCTGGCAGTGGTGAAGAACCAACCGTCTCCCATGATCTCTGAA 528
Qy 161 AlaSerIleThrThrAspIleValThrGluHisAlaAsnGlnAlaLysGluThrLeu 180
Db 529 GCTGGGCTGTCTTCAACTCATCCAGACAGACAGAACCCAGCCCGCGGAGCTGGTC 588
Qy 181 TyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGlyMetPhe 200
Db 589 CAGGGCTGAGCCTGAGTGAGTGGATGTCATCGTCACGGTCTCGGAGAGCGGGCTGCTC 648
Qy 201 SerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAspGlnAsn 220
Db 649 CATGAGGTGCTGAACGGCTCTAGATGCC-----GTGGGATCTCCCTCGCGGCTCG 684
Qy 221 HisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAlaGlySer 240
Db 685 TGGGAGGAGCTGTGAAGATGCT-----GTGGGATCTCCCTCGCGGCTCG 732
Qy 241 ThrAspCysValCysTyrSer-----ThrValGlyThr 251
Db 733 GCACAGCGCTGGCGGAGCAGTGAACACGACGCGGGGATTTGACCCAGCCCTGGGCTC 792
Qy 252 SerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspVal 271
Db 792 SerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspVal 271
```

547	ATGTCCTA-GTCTCTCTGGATGAGGTCTCGGGCTGTGTGCACCCCTGTGCAGAGCGGTAGCCCC	605
Qy	GlucIthrTrpAspValHis- ::: :::	72
Db	GAGCACA ::: :::	665
Qy	GlyLysTrpGlnLysMetGluLysProTyrAlaPheThrValHisCysValLysArgAla	92
Db	::: ::: :::	725
Qy	ArgArgHisArgTyrLysTrpAlaGlnValThrPheTrpCysProGluGluGlnLeuCys	112
Db	::: ::: :::	767
Qy	HisLeuTrpLeuGlnThrLeuArgGluMetLeuGluLysLeuThrSerArgProLysHis	132
Db	::: ::: :::	822
Qy	LeuLeuValPheIleAsnProPheGlyGlyGlnGlyLysArgIleTyrGluArg	152
Db	::: ::: :::	887
Qy	LysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHis	172
Db	::: ::: :::	947
Qy	AlaAsnGlnAlaySGLuThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleVal	192
Db	::: ::: :::	1007
Qy	CysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGln	212
Db	::: ::: :::	1061
Qy	ArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArg	232
Db	::: ::: :::	1091
Qy	IleGlyIleIleProAlaglySerThrAspCysValCys-	245
Db	::: ::: :::	1151
Qy	TyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValVal	263
Db	::: ::: :::	1211
Qy	GlyAspSerLeuAlaMetAspValSerSerValHisAsnSerThrLeuLeuArgTyr	283
Db	::: ::: :::	1271
Qy	SerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGluLysLys	303
Db	::: ::: :::	1331
Qy	ArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHis	323
Db	::: ::: :::	1391
Qy	CysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArg-As	343
Db	::: ::: :::	1430
Qy	pArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGln	359
Db	::: ::: :::	1479

```

; APPLICANT: Pu, XIA
; APPLICANT: Richard, D'ANDREA J
; APPLICANT: Jennifer, BAMBLE R
; APPLICANT: Mathew, VADAS A
; TITLE OF INVENTION: SPHINGOSINE KINASE ENZYME
; FILE REFERENCE: PITSON=1
; CURRENT APPLICATION NUMBER: US/09/959,897
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: PCT/AU00/00457
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: AU PQ 0339
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: AU PQ 1504
; PRIOR FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1205
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (33)..(1184)
; OTHER INFORMATION:
US-09-959-897-1

Alignment Scores:
Pred. No.: 4,98e-20 Length: 1205
Score: 266.50 Matches: 107
Percent Similarity: 41.37% Conservative: 68
Best Local Similarity: 25.30% Mismatches: 169
Query Match: 9.23% Indels: 79
DB: 4 Gaps: 16

US-10-631-958-10 (1-537) x US-09-959-897-1 (1-1205)

Qy 129 ArgProlyHisLeuLeuValPheIleAasnProPheGlyLysGlyGlnGlyLysArg 148
Db 69 CGGCGCTGCGCGTGTGCTGCTGAACCGCGCGCGCAAGGCGCAAGCCCTTGCGAG 128
Qy 149 IleTyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrThrAspIleIle 168
Db 129 CTCCTTCGCGAGTCACGTGCGACGCCCTTTTGGCTGAGGCTGAAATCTCTTTCACGCTGATG 188
Qy 169 ValThrGluHisAlaAasnGlnAlaLysGluThrLeuTyrGluIleAasnIleAspLysTyr 188
Db 189 CTCACCTAGCGCGCGGAACCAACGCGCGGAGCTGTGCGTTCGAGGAGCTGGCGCGCTGG 248
Qy 189 AspGlyIleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIle 208
Db 249 GACGCTCTGGTGGTCATGTCTGGAGACGGGCTGATGCACGAGGTGGTCAACGGGCTCATG 308
Qy 209 GlyArgThrGlnArgSerAlaGlyValAspGlnAasnHisProArgAlaValLeuValPro 228
Db 309 GAGCGGCTGACTGGGAGACCGCCATCCAGAAG----- 341
Qy 229 SerSerLeuArgIleGlyIleIleProAlaGlySerThrAspCysValCys----- 245
Db 342 -----CCCTGTGTAGCTCTCCACAGCGCTCTGGCAACGGCGCTGGCAGCTTCCTTG 392
Qy 246 -----TyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
Db 393 AACCATTTATGCTGGCTATGACAGGTCACCAATAGACCTCTCTGACCACATGCACGCTA 452
Qy 261 IleValValGlyAspSerLeuAla---MetAspValSerSerValHisHisAasnSerThr 279
Db 453 TTGCTGTGCGCGCGGCTGCTGTCAACCATGAACCTGTCTCTCTGCACACGCGCTTCGGGG 512
Qy 280 LeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAsp 299
Db 513 CTGCGCTCTTCTCTGTGCTAGCGCTGGCGGCTTCATWTGCTGATGTGGACCTAGAG 572
Qy 300 SerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPhe 319

```

```
Db 573 AGTGAGAGTATCGCGCTGCGGAGATGCGTTCACTCTGGGCACTTTCTCGCTG 632
Qy 320 LeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGly 339
Db 633 GCACCTTGGCAGCTTACCGCGCGCACTGGCTTACCTCCCTGTAGGAAGA---GTGGT 689
Qy 340 SerProArgAspArgLysPheValCysArgAlaGlyCysPheValCysArgGlnSerLysGln 359
Db 690 TCCAGACACCTGCTCCCC-----GTTGGTCCAGCAGGCGCCGGTA 734
Qy 360 GlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal--- 378
Db 735 GATGCACAC-----CTTGTGCCACTGGAG-----GAGCCAGTGC 770
Qy 379 GluGluTrpGlnValValCysGlyLys-----PheLeuAlaAlaAsnAlaThr 394
Db 771 TCTCACTGGACAGTGTGTCGCCGACAGGACTTTGTGTAGTCTGCTGGCACTGCTGCAC 830
Qy 395 AsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGly 414
Db 831 CACTGGGCAGTGTGCTGTCGCCACCATGCGCGCTGTGCAGCT----- 878
Qy 415 AspGlySerSerAspLeuLeuLeuLeuArgLys---CysSerArgPheAsnPheLeuArg 433
Db 879 --GGCGTCATGCTCTGCTACGTGCGCGCGAGTGTCTCGTGCCATGCTGCTGCGC 935
Qy 434 ---PheLeuLeuLeuArgHisThrAsnGlnAspGlnPheAspPheThrPheValGluVal 452
Db 936 CTCCTTCCTGGCCATGGAGAGGCGCATATGAGTATGAATGCCCTACTTGGTATAT 995
Qy 453 TyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeu 472
Db 996 GTGCCGTGCTGCGCTTCGC-----TTGAGCCCAAGAT----- 1031
Qy 473 LysGluGlyGlyLysLysArgPheGlyHisLeCysSerSerHisProSerCysCys 492
Db 1032 -----GGGAAAGGTATGTTGCA----- 1049
Qy 493 ThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaLeuGlu 512
Db 1050 -----GTGGATGGGAAATTGATGTAGCGAGCGCGTGCAG 1085
Qy 513 ValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyLleGluGluAsnPro--- 531
Db 1086 GCCAGGTGACCCAACTACTTCTGGATGTCTAGCGGTTCGTTGGAGCCCGCCGAGC 1145
Qy 532 ---LysPro 533
Db 1146 TGAAGCCC 1154

RESULT 9
US-09-970-516-1
; Sequence 1, Application US/09970516
; Patent No. 6610534
; GENERAL INFORMATION:
; APPLICANT: No. 6610534artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1155
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1155)
; OTHER INFORMATION:
US-09-970-516-1
```

```
Alignment Scores:
Pred. No.: 7,83e-20 Length: 1155
Score: 264.50 Matches: 107
Percent Similarity: 41.37% Conservatives: 68
Best Local Similarity: 25.30% Mismatches: 169
Query Match: 9.16% Indels: 79
DB: 4 Gaps: 16

US-10-631-958-10 (1-537) x US-09-970-516-1 (1-1155)
Qy 129 ArgProLysHisLeuLeuValPheIleAsnProPheGlyGlyLysGlyGlnGlyLysArg 148
Db 37 CGGCGCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 96
Qy 149 IleTyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleLeu 168
Db 97 CTCTTCGGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 156
Qy 169 ValThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyr 188
Db 157 CTACTGAGCGCGGACCAACACCGCGGAGCTGGTGGCTCGGAGGAGCTGGGCGCTGG 216
Qy 189 AspGlyLleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIle 208
Db 217 GACGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 276
Qy 209 GlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValPro 228
Db 277 GAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 309
Qy 229 SerSerLeuArgIleGlyIleIleProAlaGlySerThrAspCysValCys--- 245
Db 310 -----CCCTGTGTAGCTTCCAGAGGCTCTGGCAACGCGCTGGCAGCTTCTCTTG 360
Qy 246 -----TyrSerThrValGlyThrSerAspAlaGluThrSerAlaLeuHis 260
Db 361 AACCATTTATGCTGCTATGAGAGGTCAACATGAAGACCTCTCTGACCACTGCACGCTA 420
Qy 261 IleValValGlyAspSerLeuAla---MetAspValSerSerValHisHisAsnSerThr 279
Db 421 TTGCTGTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Qy 280 LeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleLeuLysAsp 299
Db 481 CTGCGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Qy 300 SerGluLysLysArgTrpLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPhe 319
Db 541 AGTGAGAAATATCGCGCTCTGGGCGAGATGCGCTTCACTCTGGGCACTTCTCTGCTGCTG 600
Qy 320 LeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGly 339
Db 601 GCAGCCTGCGCAGCTTACCGCGCGGAGCTGGCTTACCTCTCTGTAGGAAGA---GTGGGT 657
Qy 340 SerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGln 359
Db 658 TCCAGACACCTGCTCCCTCC-----GTTGTGGTCCAGCAGGCGCCGGTA 702
Qy 360 GlnLeuGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal--- 378
Db 703 GATGCACAC-----CTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 738
Qy 379 GluGluTrpGlnValValCysGlyLys-----PheLeuAlaAlaAsnAlaThr 394
Db 739 TCTCACTGGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 798
Qy 395 AsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGly 414
Db 799 CACTGGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 846
Qy 415 AspGlySerSerAspLeuLeuLeuLeuArgLys---CysSerArgPheAsnPheLeuArg 433
```

Db 847 ---GGCGTCATCATCTGTTCTACGTGCGGGCGGAGGTCTCTCGTCCATGCTGCTGCGC 903  
Qy 434 ---PheLeuIleArgHisThrAsnGlnAspGlnPheAspPheThrPheValGluVal 452  
Db 904 CTCCTCTGGCCATGCAGAGGCGAGCATATGAGTATGAATGATGATGATGATGATGAT 963  
Qy 453 TyrArgValLysPheGlnPheThrSerLysHisMetGluAspGluAspSerAspLeu 472  
Db 964 GTGCCCGTGTGCTTCCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 999  
Qy 473 LysGluGlyGlyLysLysArgPheGlyHisLysLysLysLysLysLysLysLysLys 492  
Db 1000 ---GGGAAGGTGTGTTTCA----- 1017  
Qy 493 ThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerProAlaIleGlu 512  
Db 1018 ---GTGGATGGGAATTCATGTTAGCGAGCGCGTGCAG 1053  
Qy 513 ValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGluGluAsnPro--- 531  
Db 1054 GGCCAGGTGCACCCAACTACTTCTGGATGTGTCAGCGTTCGTTGGAGCGCCCGCCAGC 1113  
Qy 532 ---LysPro 533  
Db 1114 TGAAGCCC 1122

## RESULT 10

US-09-949-016-1155  
; Sequence 1155, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1155  
; LENGTH: 1783  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-1155

## Alignment Scores:

Pred. No.:	2,15e-19	Length:	1783
Score:	263.50	Matches:	107
Percent Similarity:	41.22%	Conservative:	69
Best Local Similarity:	25.06%	Mismatches:	164
Query Match:	9.12%	Indels:	87
DB:	4	Gaps:	17

US-10-631-958-10 (1-537) x US-09-949-016-1155 (1-1783)

Qy 129 ArgProLysHisLeuValPheIleAsnProPheGlyLysGlyGlnGlyLysArg 148  
Db 373 CGGCCCTGCGCGT 432  
Qy 149 IleTyrGluArgLysValAlaProLeuPheThrLeuAlaSerIleThrAspIleIle 168  
Db 433 CTCCTCCGAGTACGTGCGAGCCCTTTTGGCTGAGGTGAATCTCTTCACGCTGATG 492  
Qy 169 ValThrGluHisAlaAsnGlnAlaLysGluThrLeuTyrGluIleAsnIleAspLysTyr 188  
Db 493 CTCACGTAGCGCGGAACACACGCGGAGCTGGTGTGGAGGAGCTGGCGCGCTGG 552

Qy 189 AspGlyIleValCysValGlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIle 208  
Db 553 GACCTCTGTGGTGCATCTCTGGAGACGGGCTGATGCACAGGTTGGTGAACGGGCTCATG 612  
Qy 209 GlyArgThrGlnArgSerAlaGlyValAspGlnAsnHisProArgAlaValLeuValPro 228  
Db 613 GAGCGGCTGACTGGGAGACCGCATCCAGAG----- 645  
Qy 229 SerSerLeuArgIleGlyIleIleProAlaGlySerThrAspCysValCys----- 245  
Db 646 ---CCCTGTGTAGCTCCAGACGGCTCTGGCAACGGCTGGCAGCTTCCTCTG 696  
Qy 246 ---TyrSerThrValClyThrSerAspAlaGluThrSerAlaLeuHis 260  
Db 697 AACCATATGCTGCTATGACAGGTCAACATGAAGACCTCTCGACCACTGCACGCTA 756  
Qy 261 IleValValGlyAspSerLeuAla---MetAspValSerSerValHisHisAsnSerThr 279  
Db 757 TTGCTGTGCCCGCGCTGCTGCACCATGACCTGCTGCTCTGCACACGGCTTCGGGG 816  
Qy 280 LeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAspIleIleLysAsp 299  
Db 817 CTGCGCTCTTCTCTGCTCAGCTGCGCTGGGCTTCATTGCTGATGTGACCTAGAG 876  
Qy 300 SerGluLysLysArgTyrPLeuGlyLeuAlaArgTyrAspPheSerGlyLeuLysThrPhe 319  
Db 877 AGTGAGAAGTATCGCGCTCTGGGGGAGATGCGCTTCACTCTGGGCACCTTCTCGCTCG 936  
Qy 320 LeuSerHisHisCysTyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGly 339  
Db 937 GCACCTCGGCACCTACCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 993  
Qy 340 SerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGlnSerLysGln 359  
Db 994 TCCAAGACACCTGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 1038  
Qy 360 GlnLeuGluGluGluGlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal--- 378  
Db 1039 GATGCACAC-----CTTGTGCCACTGGAG-----GAGCCAGTGCCTC 1074  
Qy 379 GluGluTrpGlnValValCysGlyLys-----PheLeuAlaIleAsnAlaThr 394  
Db 1075 TCTCAGCTGGACATGGTGGCGGACGAGACTTTGTGTGTAATCTCTGGCTGCTGCTGCT 1134  
Qy 395 AsnMetSerCysAlaCysArgArgSerProArgGlyLeuSerProAlaAlaHisLeuGly 414  
Db 1135 CACCTGGGCGAGTGTGTTGCTGCACCCATGGCGCTGTGTCAGCT----- 1182  
Qy 415 AspGlySerSerAspLeuIleLeuIleArgLys---CysSerArgPheAsnPheLeuArg 433  
Db 1183 ---GGCGTCATGCTCTGTTCTACGTGGGCGGAGTGTCTCGTGCATCTGCTGCTGCTG 1239  
Qy 434 PheLeuIle-----ArgHisThrAsnGlnAspGlnPheAspPheThr 448  
Db 1240 TTCTTCTGCGCATGGAGAGGCGGCGCATATG-----GAGTATGATGATGATGATG 1287  
Qy 449 PheValGluValTyrArgValLysLysPheGlnPheThrSerLysHisMetGluAspGlu 468  
Db 1288 TACTTGTATATGTCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1332  
Qy 469 AspSerAspLeuLysGluGlyLysLysArgPheGlyHisLysLysCysSerSerHisPro 488  
Db 1333 GAT-----GGGAAAGGTGTGTTTGA----- 1353  
Qy 489 SerCysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSer 508  
Db 1354 ---GTGGATGGGAAATGATGTTAGC----- 1377  
Qy 509 ProAlaIleGluValArgValHisCysGlnLeuValArgLeuPheAlaArgGlyIleGlu 528  
Db 1378 GAGCGCTGCGAGCGGAGTGCACCAACTACTTCTGATGCTGCTGCTGCTGCTGCTGCTG 1437  
Qy 529 GluAsnPro-----LysPro 533







```
QY 337 ThrValGlySerProArgAspArgLysProCysArgAlaGlyCysPheValCysArgGln 356
Db 643 ---GTAGAACTGTGGCCCTCTAAGAGACCCGCC-----
QY 357 SerLysGlnGlnLeuGluGluGlnLysAlaLeuTyrGlyLeuGluAlaGlu 376
Db 673 ---TCTACACTGGTGCAGAAAGGCCCGCCGTCACACACACACTTGTCTCTCGAGGAG 726
QY 377 AspVal---GluGluTyrGlnValValCysGlyLys-----PheLeuAlaIle 391
Db 727 CCAGTGCCTCTCATGTGACGTGTGTGACCAAGAGACTTTGTCTCGTGTGCTG 786
QY 392 AsnAlaThrAsnMetSerCysAlaCysArgSerProArgGlyLeuSerProAlaAla 411
Db 787 CTACACACCCACCTGAGCTCCGAGCTGTTTCAGCACCCATGGCGCGCTGTGAGGCT--- 843
QY 412 HisLeuGlyAspGlySerSerAspLeuIleLeuIleArgLys---CysSerArgPheAsn 430
Db 844 -----GGTGTATTGCATCTGTTCTACGTACGTGCGGGGGTGTCAAGGGCTGCG 891
QY 431 PheLeuArg---PheLeuIleArgHisThrAsnGlnGlnAspGlnPheAspPheThrPhe 449
Db 892 CTGCTGCCCTCTCTCTCGCCATGCAGAAAGGCGAAGCATATGGAACCTTGACTGTCCATAC 951
QY 450 ValGluValTyrArgValLysPheGlnPheThrSerLysHisMetGluAspGluAsp 469
Db 952 CTGGTTCATGTCGCGTGTGTTCTTCGCTGGAGCCCGAGGAGCGGCGTGT 1011
QY 470 SerAspLeuLysGluGlyLysLysArgPheGlyHisIleCysSerSerHisProSer 489
Db 1012 TCT-----GTGGATGGAGGCTGATGG--TATGTGAAGCTGTGCAGGGCCAAG----- 1057
QY 490 CysCysCysThrValSerAsnSerSerTrpAsnCysAspGlyGluValLeuHisSerPro 509
Db 1058 -----TGCACCCAAACTACCTTTGGATGGTCTGTGGCAGCAGAGATGCCCAT---CCG 1108
QY 510 Ala 510
Db 1109 GCC 1111

RESULT 13
US-09-248-796A-1756
; Sequence 1756, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 1756
; LENGTH: 1611
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-1756

Alignment Scores:
Pred. No.: 2,04e-15 Length: 1611
Score: 228.00 Matches: 81
Percent Similarity: 42.86% Conservative: 57
Best Local Similarity: 25.16% Mismatches: 143
Query Match: 7.89% Indels: 42
DB: 4 Gaps: 9

US-10-631-958-10 (1-537) x US-09-248-796A-1756 (1-1611)
QY 119 LeuArgGluMetLeuGluLysLeuThrSerArgProLysHisLeuLeuValPheIleAsn 138
```

```
Db 454 TTAACAGAACTCATATGATGGAAAAAGTATAATTTCGGCCCTCT---ATTTTGGTATTGATTAAC 510
QY 139 ProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaProLeuPhe 158
Db 511 CCGCATGGCGGCCAAGACACGCCAACTATTATTACAAATAATAAATCTTTACCAATAATTA 570
QY 159 ThrLeuAlaSerIleThrThrAspIleValThrGluHisAlaAsnGlnAlaLysGlu 178
Db 571 CARGCGCTCGTCTAATGTACGTATTTTGAACCTAAATATCATATGACACGCCACTGAG 630
QY 179 ThrLeuTyrGluLeuAsnIleAspLysTyrAspGlyIleValCysValGlyGlyAspGly 198
Db 631 ATTGCGGTGAGCTAGATGCAATGATTATGATATAATTTGTTGTTGTTGCGGATGGG 690
QY 199 MetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGlyValAsp 218
Db 691 ATACCTCATGAAGTTATCAATGGCTTCTATCTTCCTCCAGAT----- 732
QY 219 GlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIleIleProAla 238
Db 733 -----AAAGTTTACTGGCATTCACAAAAATTCAGTACTCAATTACCTTGT 780
QY 239 GlySerThrAspCysValCysTyrSerThrValGlyThrSerAspAlaGluThrSerAla 258
Db 781 GGGTCAGGGAACCGGTGTGAGTTGAGTACACATGTTAGTAAATAATGCTTCAGTTGCAACT 840
QY 259 LeuHisIleValValGlyAspSerLeuAlaMetAspValSerSerValHisHis----- 276
Db 841 CTTTATATGTTGAAGCTCATAGACAAATAATGGATTTGATGGCTATTACCCAAAGGTACA 900
QY 277 ---AsnSerThrLeuLeuArgTyrSerValSerLeuLeuGlyTyrGlyPheTyrGlyAsp 295
Db 901 GGAAGTCAAAAGATAACGAAATTTGTCATTTTAAAGTCAGTCTACGGTATTATTGCTGAT 960
QY 296 IleIleLysAspSerGluLysLysArgTrpLeuAlaArgTrpLeuAlaArgTyrAsp--- 312
Db 961 TCTGATATTGGAACAGAACATTTACGTTGGTGGTCTCTATACGGTTTGAACCTGGAGTG 1020
QY 313 -----PheSerGlyLeuLys-----ThrPheLeuSerHisHisCys 324
Db 1021 ATACAAAAGTGTCTCGGGGGCAAAATATCTTTGTGACCTATTGTGAAATACAAG--- 1077
QY 325 TyrGluGlyThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArg 344
Db 1078 TAGATAATAATTTCAGAGATTTTGAATCATGTAATGATTACTTGTAGTAATAATGAT--- 1134
QY 345 LysProCysArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLeuGluGlu 364
Db 1135 -----ACTGAAAACGAGTTACCCATTGTCACTGAA 1164
QY 365 GlnLysLysAlaLeuTyrGlyLeuGluAlaAlaGluAspVal---GluGluTrpGln-Va 383
Db 1165 GAGAACTTGCAAAATAACGACCCCTGATTTAGATCAACCTGTCCCAATGATGAGACAT 1224
QY 383 lValCysGlyLysPheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgGSe 403
Db 1225 ATTCTCAAGAAATTTCTCATTAATTGAAACATCTTGTATGTTGGCAAAATGCCATTGCTG 1284
QY 403 rProArg-----GlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuIle 421
Db 1285 TCTGCAGATACTCAGTTTTC-CCTGCTGCTCTACCGAACGATGGTTCAATGGACATGAT 1343
QY 421 eLeu 422
Db 1344 TGTC 1347

RESULT 14
US-09-614-221A-399
; Sequence 399, Application US/09614221A
; Patent No. 6723837
; GENERAL INFORMATION:
; APPLICANT: Karunanandaa, Balasulojini
```

```

; APPLICANT: Yu, Jaebuyuk
; APPLICANT: Kishore, Ganesh M.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED
; TITLE OF INVENTION: WITH STEROL SYNTHESIS AND METABOLISM
; FILE REFERENCE: 16516.075
; CURRENT APPLICATION NUMBER: US/09/614,221A
; PRIORITY FILING DATE: 2000-07-12
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 626
; SEQ ID NO 399
; LENGTH: 1875
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-614-221A-399

Alignment Scores:
Pred. No.: 2,64e-13 Length: 1875
Score: 210.50 Matches: 66
Percent Similarity: 44.62% Conservativity: 46
Best Local Similarity: 26.23% Mismatches: 102
Query Match: 7.29% Indels: 37
DB: 4 Gaps: 6

US-10-631-958-10 (1-537) x US-09-614-221A-399 (1-1875)
QY 118 ThrLeuArgGluMetLeuGluLysLeu-----ThrSerArgProLysHisLeuLeu 134
DB 631 ACTGTGGGAAGAAATCTCGGAAAAAGTTATGAAAAATTCACAGAAACAGATCGATATTA 690
QY 135 ValPheIleAsnProPheGlyGlyGlyGlnGlyLeuArgGlyLeuVal 154
DB 691 GTCATTAATTAATCCACCGTGGTAAGGTACTGCTAAAAATTTAATCTGACAAAAGCA 750
QY 155 AlaProPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsn 174
DB 751 AGGCCAATACATAGTGAAAGTGGCTGCAAAATAGAAATTCACATACAAATATGCCCGT 810
QY 175 GlnAlaLysGluThrLeuThrGluLeuIleAsnIleAspLysThrValGlyLeuVal 194
DB 811 CACGCCATCATGATATTCGCAAGATTTAGATATACGAAATACGATACCAATTCATGTGCC 870
QY 195 GlyGlyAspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSer 214
DB 871 TCGGTGTGATGATTCATACAAAGTAATTAATGGCTTTATAGAAAGCCCGACAGA--- 927
QY 215 AlaGlyValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGly 234
DB 928 -----GTGGATCGCTTCAATAAATACCGTAACTCAG----- 960
QY 235 IleIleProAlaGlySerThrAspCysValCysThrValGlyThrSerAspAla 254
DB 961 ---CTACCTTGGGTTTCAGGAATGCTATGACCAATTCATGTCATTGACAAATAACCCA 1017
QY 255 GluThrSerAlaLeuHisIleValValGlyAspSerLeuAlaMetAspVal----- 271
DB 1018 TCGTACGCGCTGCTGCTTGTCTCAAAATCCATTTGAAACAAAGAAATAGACTTAATGTGTGT 1077
QY 272 SerSerValHisAsnSerThrLeuLeuArgTyrSerValSerLeuGlyTyrGly 291
DB 1078 TCCCAACCTCTTATATGACGAATGGCAAGATTAATCTTTTGGATGACAGCTACCGC 1137
QY 292 PheTyrGlyAspIleIleLysAspSerGluLysLysArgTyrLeuGlyLeuAlaArgTyr 311
DB 1138 GTTATTGCGAATCTGATATTAACACTGAATTCATCATGATGGTCCCGTTAGGTTT 1197
QY 312 AspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGlyThrValSerPhe 331
DB 1198 AATTTGGGT-----GTAGCATTCACATTAATCCAAAGT----- 1230
QY 332 LeuProAlaGlnHisThrValGlySerProArgAspArgLysProCysArgAlaGlyCys 351
DB 1231 -----AAGAAATATCCCTCGAAGTTTTCGTC 1257
```

```

QY 352 PheValCysArgGlnSerLysGlnGlnLeuGlu 362
DB 1258 AATATGCTGCCAAATCTAAAAAGGAATAAAA 1290
```

## RESULT 15

```

US-09-949-016-3940
; Sequence 3940, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3940
; LENGTH: 1764
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3940
```

```

Alignment Scores:
Pred. No.: 1.5e-12 Length: 1764
Score: 203.50 Matches: 101
Percent Similarity: 40.48% Conservativity: 67
Best Local Similarity: 24.34% Mismatches: 167
Query Match: 7.05% Indels: 81
DB: 4 Gaps: 17

US-10-631-958-10 (1-537) x US-09-949-016-3940 (1-1764)
```

```

QY 137 IleAsnProPheGlyGlyLysGlyGlnGlyLysArgIleTyrGluArgLysValAlaPro 156
DB 397 CTGAACCCG---CGAAGCAGGCGCAAGCTT-CAGCTCTCCGAGTACGTGCAGGCC 452
QY 157 LeuPheThrLeuAlaSerIleThrThrAspIleIleValThrGluHisAlaAsnGlnAla 176
DB 453 CTTTGTGGCTGAGGCTGAAATCTCTTCACGCTGATGCTCACTGAGCGCGCAACCCACGCG 512
QY 177 LysGluThrLeuTyrGluIleAsnIleAspLysTyrAspGlyIleValCysValGlyGly 196
DB 513 CGAGACTGGATGCGATCGGAGAACCTGGCGCGCTAGGACGCTCTGGTGTCTATGCTGGA 572
QY 197 AspGlyMetPheSerGluValLeuHisGlyLeuIleGlyArgThrGlnArgSerAlaGly 216
DB 573 GACGGCTGATGTCAGAGGTGTGACGGCTCATGAGCGGCTGACTGGGAGCCGCC 632
QY 217 ValAspGlnAsnHisProArgAlaValLeuValProSerSerLeuArgIleGlyIle 236
DB 633 ATCCAGAAG-----CCCTGTGTAGCCTC 656
QY 237 ProAlaGlySerThrAspCysValCys-----TyrSerThr 248
DB 657 CCAGCAGGCTCTGGCAACGCGCTGCGAGCTTCTTGAACCATTAATATGCTGGCTATGAGCAG 716
QY 249 ValGlyThrSerAspAlaGluThrSerAlaLeuHisIleValValGlyAspSerLeuAla 268
DB 717 GTACCAATGAAGACCTCTGACCACTGACGCTATTTGCTGCGCGCGCTCTGTGCA 776
QY 269 ---MetAspValSerValHisAsnSerThrLeuLeuArgTyrSerValSerLeu 287
DB 777 CCCATGAACCTGCTCTGTCACACGCGTTTCGGGGCTGCGCCTCTTCTCTGTGCTCAGC 836
QY 288 LeuGlyTyrGlyPheTyrGlyAspIleIleLysAspSerGluLysLysArgTyrLeuGly 307
```

```
Db      |||  ::||| |||  ::|||  ::|||  ::|||  ::|||  |||  |||  |||  |||
837 CTGGCTGGGCTTCATTGCTGATGTGGACCTAGAGAGTGAGAAGTATCGGCGTCTGGGG 896
Qy      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
308 LeuAlaArgTyrAspPheSerGlyLeuLysThrPheLeuSerHisHisCysTyrGluGly 327
Db      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
897 GAGATGGCTTCACCTCTGGGCACCTCTCTGGCTGGCAGCCCTGGCACCCTACCGCGGC 956
Qy      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
328 ThrValSerPheLeuProAlaGlnHisThrValGlySerProArgAspArgLysProCys 347
Db      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
957 CGACTGGCCTACCTCCCTGTAGGAAGA--GTGGGTTCCAAGACACCTGCCTCCCCC-- 1010
Qy      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
348 ArgAlaGlyCysPheValCysArgGlnSerLysGlnGlnLeuGluGlnLysLys 367
Db      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1011 -----GTTGTGTCAGCAGGCGCCGCTAGATGCACAC----- 1043
Qy      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
368 AlaLeuTyrGlyLeuGluAlaAlaValVal---GluGluTyrGlnValValCysGly 386
Db      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1044 ---CTTGTGCCACTGGAG-----GAGCCAGTGCCTCTCACTGGACAGTGTGCCGAC 1094
Qy      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
387 Lys-----PheLeuAlaIleAsnAlaThrAsnMetSerCysAlaCysArgArg 402
Db      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1095 GAGGACITTTGTAGTCTGGCACTGCTGCACCTGCACCTGGGCGAGTGAGATGTTTGTCT 1154
Qy      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
403 SerProArgGlyLeuSerProAlaAlaHisLeuGlyAspGlySerSerAspLeuLeu 422
Db      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1155 GCACCCATGGGCGCTGTGCAGCT-----GGCGTCATGCATCTGTCTTCTAC 1199
Qy      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
423 IleArgLys---CysSerArgPheAsnPheLeuArg---PheLeuIleArgHisThrAsn 440
Db      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1200 GTGGCGGCGGAGTGTCTCGTGCCATGCTGTGGCCTCTTCCTGGCCATGGAGAAGGCG 1259
Qy      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
441 GlnGlnAspGlnPheAspPheThrPheValGluValTyrArgValLysLysPheGlnPhe 460
Db      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1260 AGGCATATGGAGTATGAATGCCCTACTTGTGTATATGTGCCCGTGGTCCCTTCGCG-- 1316
Qy      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
461 ThrSerLysHisMetGluAspGluAspSerAspLeuLysGluGlyGlyLysLysArgPhe 480
Db      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1317 -----TTGAGCCCAAGGAT-----GGGAAAGGTGTGTTT 1346
Qy      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
481 GlyHisIleCysSerSerHisProSerCysCysThrValSerAsnSerSerTrpAsn 500
Db      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1347 GCA----- 1349
Qy      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
501 CysAspGlyGluValLeuHisSerProAlaIleGluValArgValHisCysGlnLeuVal 520
Db      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1350 GTGGATGGGGAATTGATGTTAGCGAGGCCGCTGCAGGCGCCAGGTGCACCCCAACTACTTC 1409
Qy      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
521 ArgLeuPheAlaArgGlyIleGluGluAsnPro-----LysPro 533
Db      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1410 TGGATGTCAGCGGTTGCGTGAGCCCCGCCAGCTGGAAAGCCC 1454
```

Search completed: September 6, 2005, 02:34:14  
Job time : 349.735 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model  
Run on: September 3, 2005, 03:56:52 ; Search time 347.06 Seconds  
(without alignments)  
1891.374 Million cell updates/sec

Title: US-10-631-958-11  
Perfect score: 3025  
Sequence: 1 HEANGAPLGVRRAPPWRT.....OLVRLFARGIENPKPDSHS 562

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 6959266 seqs, 1168006243 residues

Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:\*

1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pcp.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pcp.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pcp.\*  
4: /cgn2\_6/ptodata/1/paa/US080\_COMB.pcp.\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pcp.\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pcp.\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pcp.\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pcp.\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pcp.\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pcp.\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pcp.\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pcp.\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pcp.\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pcp.\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pcp.\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pcp.\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pcp.\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pcp.\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pcp.\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pcp.\*  
21: /cgn2\_6/ptodata/1/paa/US097A\_COMB.pcp.\*  
22: /cgn2\_6/ptodata/1/paa/US097B\_COMB.pcp.\*  
23: /cgn2\_6/ptodata/1/paa/US098\_COMB.pcp.\*  
24: /cgn2\_6/ptodata/1/paa/US099A\_COMB.pcp.\*  
25: /cgn2\_6/ptodata/1/paa/US099B\_COMB.pcp.\*  
26: /cgn2\_6/ptodata/1/paa/US100\_COMB.pcp.\*  
27: /cgn2\_6/ptodata/1/paa/US101\_COMB.pcp.\*  
28: /cgn2\_6/ptodata/1/paa/US102\_COMB.pcp.\*  
29: /cgn2\_6/ptodata/1/paa/US103\_COMB.pcp.\*  
30: /cgn2\_6/ptodata/1/paa/US104\_COMB.pcp.\*  
31: /cgn2\_6/ptodata/1/paa/US105\_COMB.pcp.\*  
32: /cgn2\_6/ptodata/1/paa/US106\_COMB.pcp.\*  
33: /cgn2\_6/ptodata/1/paa/US107\_COMB.pcp.\*  
34: /cgn2\_6/ptodata/1/paa/US108\_COMB.pcp.\*  
35: /cgn2\_6/ptodata/1/paa/US109\_COMB.pcp.\*  
36: /cgn2\_6/ptodata/1/paa/US110\_COMB.pcp.\*  
37: /cgn2\_6/ptodata/1/paa/US60\_COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3025	100.0	562	25	US-09-969-896-11
2	3025	100.0	562	32	US-10-631-958-11
3	2888	95.5	537	1	PCT-US02-24623-13
4	2888	95.5	537	1	PCT-US02-31373-40
5	2888	95.5	537	1	PCT-US03-21730-121
6	2888	95.5	537	25	US-09-969-896-10
7	2888	95.5	537	28	US-10-262-511-40
8	2888	95.5	537	28	US-10-262-511A-40
9	2888	95.5	537	30	US-10-486-306-13
10	2888	95.5	537	32	US-10-618-941-121
11	2888	95.5	537	32	US-10-631-958-10
12	2880	95.2	537	29	US-10-315-597A-2
13	2874.5	95.0	544	27	US-10-170-205E-15652
14	2588	85.6	481	1	PCT-US02-24623-12
15	2588	85.6	481	30	US-10-486-306-12
16	2456.5	81.2	471	22	US-09-784-810A-6
17	2456.5	81.2	471	22	US-09-784-810C-6
18	2388.5	79.0	485	37	US-60-243-468-1391
19	2350	77.7	536	1	PCT-US03-30720-1291
20	2215	73.2	746	1	PCT-US01-08631-43900
21	1714.5	56.7	727	1	PCT-US01-08631-43902
22	1640.5	54.2	326	22	US-09-784-810A-11
23	1640.5	54.2	326	22	US-09-784-810C-11
24	1640.5	54.2	326	25	US-09-969-896-2
25	1640.5	54.2	326	32	US-10-631-958-2
26	1361.5	45.0	454	1	PCT-US01-08631-43903
27	1159	38.3	255	1	PCT-US03-30720-2504
28	1055	34.9	228	28	US-10-218-140-3172
29	1032	34.1	190	1	PCT-US01-42950-573
30	1032	34.1	190	30	US-10-416-993-573
31	986	32.6	182	22	US-09-784-810A-29
32	986	32.6	182	22	US-09-784-810C-29
33	986	32.6	182	28	US-10-218-140-4294
34	675	22.3	596	20	US-09-614-150-35799
35	675	22.3	596	20	US-09-614-150A-35799
36	675	22.3	596	37	US-60-191-637-35410
37	664	22.0	136	1	PCT-US01-08631-43899
38	647	21.4	144	22	US-09-784-810A-8
39	647	21.4	144	22	US-09-784-810C-8
40	628	20.8	136	1	PCT-US01-08631-43901
41	611.5	20.2	588	30	US-10-425-115-361076
42	597.5	19.8	575	30	US-10-425-115-270874
43	594	19.6	532	37	US-60-474-239-5
44	504.5	16.7	496	27	US-10-170-205E-6540
45	480.5	15.9	613	30	US-10-437-963-195284

ALIGNMENTS

RESULT 1  
US-09-969-896-11  
; Sequence 11, Application US/0969896  
; GENERAL INFORMATION:  
; APPLICANT: Kossida, Sophia  
; TITLE OF INVENTION: Regulation of human Sphingosine  
; TITLE OF INVENTION: Kinase-Like Protein  
; FILE REFERENCE: 004974.00594  
; CURRENT APPLICATION NUMBER: US/09/969,896  
; CURRENT FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: US 60/238,005  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: US 60/314,113  
; PRIOR FILING DATE: 2001-08-23  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 562  
; TYPE: PRT



QY 86 BETDVHGKHGSGKQKMEKPYAFTVHCVRARRHRKWAQVTFWCPBEQLCHLWLOTLR 145  
Db 61 ESTDVHGKHGSGKQKMEKPYAFTVHCVRARRHRKWAQVTFWCPBEQLCHLWLOTLR 120  
QY 146 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTTDIIIVTEHANOAKETL 205  
Db 121 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTTDIIIVTEHANOAKETL 180  
QY 206 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPAGS 265  
Db 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPAGS 240  
QY 266 TDCVCYSTVGTSDAETSAHLIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYGYGDI IKDS 325  
Db 241 TDCVCYSTVGTSDAETSAHLIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYGYGDI IKDS 300  
QY 326 EKRWGLGARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQ 385  
Db 301 EKRWGLGARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQ 360  
QY 386 LEEQKALYGLEAAEDVEEQQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 445  
Db 361 LEEQKALYGLEAAEDVEEQQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420  
QY 446 ILIRKCSRFNRLIRHTNQDQDFTFVEVYRVKQFTSKHMEDESDLKGGKKRF 505  
Db 421 ILIRKCSRFNRLIRHTNQDQDFTFVEVYRVKQFTSKHMEDESDLKGGKKRF 480  
QY 506 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSSH 562  
Db 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSSH 537

## RESULT 4

PCT-US02-31373-40

; Sequence 40, Application PC/TUS0231373

; GENERAL INFORMATION:

; APPLICANT: Curagen Corporation  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Payman, John A.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Ju, Jingfang  
; APPLICANT: Li, Li  
; APPLICANT: Guo, Xiaojia (Sasha)  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Ort, Tatiana  
; APPLICANT: Gorman, Linda  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Zhong, Mei  
; APPLICANT: Carterton, Elina  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Stone, David J.  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shinkets, Richard A.  
; APPLICANT: Rothenberg, Mark E.  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Ages, Michele L.  
; APPLICANT: Berghs, Constance  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-462C-061  
; CURRENT APPLICATION NUMBER: PCT/US02/31373  
; CURRENT FILING DATE: 2003-09-28  
; PRIOR APPLICATION NUMBER: 10/262,511  
; PRIOR FILING DATE: 2002-10-01

; PRIOR APPLICATION NUMBER: 60/326,483  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/373,815  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/327,917  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/381,642  
; PRIOR FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: 60/328,029  
; PRIOR FILING DATE: 2002-10-09  
; PRIOR APPLICATION NUMBER: 60/381,038  
; PRIOR FILING DATE: 2002-05-16  
; PRIOR APPLICATION NUMBER: 60/328,056  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/373,260  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: 60/373,826  
; PRIOR FILING DATE: 2002-04-19  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: CuraSequelist version 0.1  
; SEQ ID NO 40  
; LENGTH: 537  
; TYPE: PRT  
; ORGANISM: CG143216-01  
PCT-US02-31373-40

Query Match 95.5%; Score 2888; DB 1; Length 537;  
Best Local Similarity 100.0%; Pred. No. 5e-279;  
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 MGATGAAPLOSVLWVQORCAVSLPARALLRWRSPGPGAGAPGADACSVPVSEIIAV 85  
Db 1 MGATGAAPLOSVLWVQORCAVSLPARALLRWRSPGPGAGAPGADACSVPVSEIIAV 60  
QY 86 BETDVHGKHGSGKQKMEKPYAFTVHCVRARRHRKWAQVTFWCPBEQLCHLWLOTLR 145  
Db 61 BETDVHGKHGSGKQKMEKPYAFTVHCVRARRHRKWAQVTFWCPBEQLCHLWLOTLR 120  
QY 146 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTTDIIIVTEHANOAKETL 205  
Db 121 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTTDIIIVTEHANOAKETL 180  
QY 206 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPAGS 265  
Db 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRVLPSSLRIGIIPAGS 240  
QY 266 TDCVCYSTVGTSDAETSAHLIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYGYGDI IKDS 325  
Db 241 TDCVCYSTVGTSDAETSAHLIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYGYGDI IKDS 300  
QY 326 EKRWGLGARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQ 385  
Db 301 EKRWGLGARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRCAGCFVCRQSKQ 360  
QY 386 LEEQKALYGLEAAEDVEEQQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 445  
Db 361 LEEQKALYGLEAAEDVEEQQVCGKFLAINATNMSCACRRSPRGLSPAHLGDGSSDL 420  
QY 446 ILIRKCSRFNRLIRHTNQDQDFTFVEVYRVKQFTSKHMEDESDLKGGKKRF 505  
Db 421 ILIRKCSRFNRLIRHTNQDQDFTFVEVYRVKQFTSKHMEDESDLKGGKKRF 480  
QY 506 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSSH 562  
Db 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIEENPKPDSSH 537

## RESULT 5

PCT-US03-21730-121

; Sequence 121, Application PC/TUS0321730

; GENERAL INFORMATION:

; APPLICANT: SUGEN, INC.



```
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536/0461
; CURRENT APPLICATION NUMBER: PCT/US03/21730
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 121
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-21730-121

Query Match 95.5%; Score 2888; DB 1; Length 537;
Best Local Similarity 100.0%; Pred. No. 5e-279;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 MGATGAEPLOSVLWKQORCAVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV 85
Db 1 MGATGAEPLOSVLWKQORCAVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV 60
Qy 86 EETDVHGKHQSGKWQMEKPYAFTVHCVKRARRHWKWAQVTFWCPPEEQCHLWLQTLR 145
Db 61 EETDVHGKHQSGKWQMEKPYAFTVHCVKRARRHWKWAQVTFWCPPEEQCHLWLQTLR 120
Qy 146 EMLEKLTSRPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 205
Db 121 EMLEKLTSRPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180
Qy 206 YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGS 265
Db 181 YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGS 240
Qy 266 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS 325
Db 241 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS 300
Qy 326 EKKRWGLGARYDFSLGKTLFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCROSKQ 385
Db 301 EKKRWGLGARYDFSLGKTLFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCROSKQ 360
Qy 386 LEEQKALYGLEAAEDVEEVQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 445
Db 361 LEEQKALYGLEAAEDVEEVQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 420
Qy 446 ILIRKSRFNLRLIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGGKKRF 505
Db 421 ILIRKSRFNLRLIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGGKKRF 480
Qy 506 GHICSSHPSCCTVSNSNWNCDGEVLHSPAIEVRVHCOLVRLFARGIEENPKPDHS 562
Db 481 GHICSSHPSCCTVSNSNWNCDGEVLHSPAIEVRVHCOLVRLFARGIEENPKPDHS 537

RESULT 6
US-09-969-896-10
; Sequence 10, Application US/09969896
; GENERAL INFORMATION:
; APPLICANT: Smithson, Glenn
; APPLICANT: Miller, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.

; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536/0461
; CURRENT APPLICATION NUMBER: PCT/US03/21730
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 121
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-21730-121

Query Match 95.5%; Score 2888; DB 1; Length 537;
Best Local Similarity 100.0%; Pred. No. 5e-279;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 MGATGAEPLOSVLWKQORCAVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV 85
Db 1 MGATGAEPLOSVLWKQORCAVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV 60
Qy 86 EETDVHGKHQSGKWQMEKPYAFTVHCVKRARRHWKWAQVTFWCPPEEQCHLWLQTLR 145
Db 61 EETDVHGKHQSGKWQMEKPYAFTVHCVKRARRHWKWAQVTFWCPPEEQCHLWLQTLR 120
Qy 146 EMLEKLTSRPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 205
Db 121 EMLEKLTSRPKHLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180
Qy 206 YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGS 265
Db 181 YEINIDKYDGIIVCGDGMFSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGIIPAGS 240
Qy 266 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS 325
Db 241 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGYGDIIKDS 300
Qy 326 EKKRWGLGARYDFSLGKTLFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCROSKQ 385
Db 301 EKKRWGLGARYDFSLGKTLFLSHHCYEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCROSKQ 360
Qy 386 LEEQKALYGLEAAEDVEEVQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 445
Db 361 LEEQKALYGLEAAEDVEEVQVCGKFLAINATNMSCACRRSPRGLSPAHLGDSDDL 420
Qy 446 ILIRKSRFNLRLIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGGKKRF 505
Db 421 ILIRKSRFNLRLIRHTNQDQDFTFVEVYRVKKFQFTSKHMEDESDLKEGGKKRF 480
Qy 506 GHICSSHPSCCTVSNSNWNCDGEVLHSPAIEVRVHCOLVRLFARGIEENPKPDHS 562
Db 481 GHICSSHPSCCTVSNSNWNCDGEVLHSPAIEVRVHCOLVRLFARGIEENPKPDHS 537

RESULT 6
US-09-969-896-10
; Sequence 10, Application US/09969896
; GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 537
```



```
Best Local Similarity 100.0%; Pred. No. 5e-279;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 MGATGAAPLQSVLWVKQORCAVSLPARALLRWRSPPGAGAGADACSVPVSEIIAV 85
Db 1 MGATGAAPLQSVLWVKQORCAVSLPARALLRWRSPPGAGAGADACSVPVSEIIAV 60
Qy 86 EETDVHGKHQSGKWKQKMEKEPYAFTVHCVKRARRHRWKWAQVTFWCPEEQIHLWLQTLR 145
Db 61 EETDVHGKHQSGKWKQKMEKEPYAFTVHCVKRARRHRWKWAQVTFWCPEEQIHLWLQTLR 120
Qy 146 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 205
Db 121 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180
Qy 206 YEINIDKYDGIIVCGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 265
Db 181 YEINIDKYDGIIVCGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 240
Qy 266 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGFYGDIIKDS 325
Db 61 EETDVHGKHQSGKWKQKMEKEPYAFTVHCVKRARRHRWKWAQVTFWCPEEQIHLWLQTLR 120
Qy 146 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 205
Db 121 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180
Qy 206 YEINIDKYDGIIVCGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 265
Db 181 YEINIDKYDGIIVCGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 240
Qy 266 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGFYGDIIKDS 325
Db 61 EETDVHGKHQSGKWKQKMEKEPYAFTVHCVKRARRHRWKWAQVTFWCPEEQIHLWLQTLR 120
Qy 146 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 205
Db 121 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180
Qy 206 YEINIDKYDGIIVCGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 265
Db 181 YEINIDKYDGIIVCGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 240
Qy 266 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGFYGDIIKDS 325
Db 61 EETDVHGKHQSGKWKQKMEKEPYAFTVHCVKRARRHRWKWAQVTFWCPEEQIHLWLQTLR 120

RESULT 9
US-10-486-306-13
; Sequence 13, Application US/10486306
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SPKS AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-104C-US
; CURRENT APPLICATION NUMBER: US/10/486,306
; PRIOR FILING DATE: 2004-02-05
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/310,362
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,501
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-486-306-13

Query Match 95.5%; Score 2888; DB 30; Length 537;
Best Local Similarity 100.0%; Pred. No. 5e-279;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 MGATGAAPLQSVLWVKQORCAVSLPARALLRWRSPPGAGAGADACSVPVSEIIAV 85
Db 1 MGATGAAPLQSVLWVKQORCAVSLPARALLRWRSPPGAGAGADACSVPVSEIIAV 60
Qy 86 EETDVHGKHQSGKWKQKMEKEPYAFTVHCVKRARRHRWKWAQVTFWCPEEQIHLWLQTLR 145
Db 61 EETDVHGKHQSGKWKQKMEKEPYAFTVHCVKRARRHRWKWAQVTFWCPEEQIHLWLQTLR 120

RESULT 10
US-10-618-941-121
; Sequence 121, Application US/10618941
; GENERAL INFORMATION:
; APPLICANT: WHITE, DAVID
; APPLICANT: MANNING, GERARD
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618,941
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 121
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-618-941-121

Query Match 95.5%; Score 2888; DB 32; Length 537;
Best Local Similarity 100.0%; Pred. No. 5e-279;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 MGATGAAPLQSVLWVKQORCAVSLPARALLRWRSPPGAGAGADACSVPVSEIIAV 85
Db 1 MGATGAAPLQSVLWVKQORCAVSLPARALLRWRSPPGAGAGADACSVPVSEIIAV 60
Qy 86 EETDVHGKHQSGKWKQKMEKEPYAFTVHCVKRARRHRWKWAQVTFWCPEEQIHLWLQTLR 145
Db 61 EETDVHGKHQSGKWKQKMEKEPYAFTVHCVKRARRHRWKWAQVTFWCPEEQIHLWLQTLR 120
Qy 146 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 205
Db 121 EMLEKLTSRPKHLLVFINPFGKGQKRIYERKVAFLFTLASITTDIIIVTEHANOAKETL 180
Qy 206 YEINIDKYDGIIVCGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 265
Db 181 YEINIDKYDGIIVCGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 240
Qy 266 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGFYGDIIKDS 325
Db 241 TDCVCYSTVGTSDAETSALHIVVGDLSLAMDVSSVHHNSTLLRYSVSLGFGFYGDIIKDS 300
```

QY 326 EKKRWGLGARYDFSLGKTLFSLHHCEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCRQSKQ 385  
Db 301 EKKRWGLGARYDFSLGKTLFSLHHCEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCRQSKQ 360  
QY 386 LEEEOKKALYGLEAAEDVEEQVCGKFLAINATNNSCACRRSPRGLSPAHLGDSDDL 445  
Db 361 LEEEOKKALYGLEAAEDVEEQVCGKFLAINATNNSCACRRSPRGLSPAHLGDSDDL 420  
QY 446 ILIRKCSRNFRLIRHTNQDQDFTFVEVYRVKQFTSKHMEDESDLKEGGKKRF 505  
Db 421 ILIRKCSRNFRLIRHTNQDQDFTFVEVYRVKQFTSKHMEDESDLKEGGKKRF 480  
QY 506 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 562  
Db 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 537

RESULT 11  
US-10-631-958-10  
; Sequence 10, Application US/10631958  
; GENERAL INFORMATION:  
; APPLICANT: Kossida, Sophia  
; TITLE OF INVENTION: Regulation of human Sphingosine  
; FILE REFERENCE: 004974.00594  
; CURRENT APPLICATION NUMBER: US/10/631,958  
; PRIOR FILING DATE: 2003-08-01  
; PRIOR APPLICATION NUMBER: US/09/969,896  
; PRIOR FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: US 60/238,005  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: US 60/314,113  
; PRIOR FILING DATE: 2001-08-23  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 537  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-631-958-10

Query Match 95.5%; Score 2888; DB 32; Length 537;  
Best Local Similarity 100.0%; Pred. No. 5e-279;  
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 26 MGATGAAPLQSVLVKQORCAVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV 85  
Db 1 MGATGAAPLQSVLVKQORCAVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV 60  
QY 86 EETDVHGKHQSGKQKMEKPYAFTVHCVRARRHRKWAQVTFWCPEEQQLCHLWQLTLR 145  
Db 61 EETDVHGKHQSGKQKMEKPYAFTVHCVRARRHRKWAQVTFWCPEEQQLCHLWQLTLR 120  
QY 146 EMLEKLTSRPKHLVFINPFGKGQGRKRIYERKVAFLTLASITTDIIIVTEHANOAKETL 205  
Db 121 EMLEKLTSRPKHLVFINPFGKGQGRKRIYERKVAFLTLASITTDIIIVTEHANOAKETL 180  
QY 206 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 265  
Db 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 240  
QY 266 TDCVCYSTVGTSDAETSAHLIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYFGYGDIIKDS 325  
Db 241 TDCVCYSTVGTSDAETSAHLIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYFGYGDIIKDS 300  
QY 326 EKKRWGLARYDFSLGKTLFSLHHCEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCRQSKQ 385  
Db 301 EKKRWGLARYDFSLGKTLFSLHHCEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCRQSKQ 360  
QY 386 LEEEOKKALYGLEAAEDVEEQVCGKFLAINATNNSCACRRSPRGLSPAHLGDSDDL 445  
Db 361 LEEEOKKALYGLEAAEDVEEQVCGKFLAINATNNSCACRRSPRGLSPAHLGDSDDL 420

QY 446 ILIRKCSRNFRLIRHTNQDQDFTFVEVYRVKQFTSKHMEDESDLKEGGKKRF 505  
Db 421 ILIRKCSRNFRLIRHTNQDQDFTFVEVYRVKQFTSKHMEDESDLKEGGKKRF 480  
QY 506 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 562  
Db 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 537

RESULT 12  
US-10-315-597A-2  
; Sequence 2 Application US/10315597A  
; GENERAL INFORMATION:  
; APPLICANT: Sugiyura, Masako  
; APPLICANT: Kono, Keita  
; APPLICANT: Kohama, Takafumi  
; TITLE OF INVENTION: Ceramide Kinase and DNA Encoding It  
; FILE REFERENCE: 02658CIP/HG  
; CURRENT APPLICATION NUMBER: US/10/315,597A  
; PRIOR FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: JP 2000-178039  
; PRIOR FILING DATE: 2000-06-14  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 537  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-315-597A-2

Query Match 95.2%; Score 2880; DB 29; Length 537;  
Best Local Similarity 99.6%; Pred. No. 3.2e-278;  
Matches 535; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 26 MGATGAAPLQSVLVKQORCAVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV 85  
Db 1 MGATGAAPLQSVLVKQORCAVSLPARALLRWRSPPGAGAPGADACSVPVSEIIAV 60  
QY 86 EETDVHGKHQSGKQKMEKPYAFTVHCVRARRHRKWAQVTFWCPEEQQLCHLWQLTLR 145  
Db 61 EETDVHGKHQSGKQKMEKPYAFTVHCVRARRHRKWAQVTFWCPEEQQLCHLWQLTLR 120  
QY 146 EMLEKLTSRPKHLVFINPFGKGQGRKRIYERKVAFLTLASITTDIIIVTEHANOAKETL 205  
Db 121 EMLEKLTSRPKHLVFINPFGKGQGRKRIYERKVAFLTLASITTDIIIVTEHANOAKETL 180  
QY 206 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 265  
Db 181 YEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLVPSSLRIGIIPAGS 240  
QY 266 TDCVCYSTVGTSDAETSAHLIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYFGYGDIIKDS 325  
Db 241 TDCVCYSTVGTSDAETSAHLIIVGDSLAMDVSSVHHNSTLLRYSVSLIGYFGYGDIIKDS 300  
QY 326 EKKRWGLARYDFSLGKTLFSLHHCEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCRQSKQ 385  
Db 301 EKKRWGLARYDFSLGKTLFSLHHCEGTVSFLPAQHTVGSPPDRKPCRCAGCFVCRQSKQ 360  
QY 386 LEEEOKKALYGLEAAEDVEEQVCGKFLAINATNNSCACRRSPRGLSPAHLGDSDDL 445  
Db 361 LEEEOKKALYGLEAAEDVEEQVCGKFLAINATNNSCACRRSPRGLSPAHLGDSDDL 420  
QY 446 ILIRKCSRNFRLIRHTNQDQDFTFVEVYRVKQFTSKHMEDESDLKEGGKKRF 505  
Db 421 ILIRKCSRNFRLIRHTNQDQDFTFVEVYRVKQFTSKHMEDESDLKEGGKKRF 480  
QY 506 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 562  
Db 481 GHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCQLVRLFARGIEENPKPDSHS 537

RESULT 13

US-10-170-205E-15652  
; Sequence 15652, Application US/10170205E  
; GENERAL INFORMATION:  
; APPLICANT: ADAMS, Mark  
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN  
; FILE REFERENCE: CL001381  
; CURRENT APPLICATION NUMBER: US/10/170,205E  
; CURRENT FILING DATE: 2002-06-13  
; NUMBER OF SEQ ID NOS: 40312  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 15652  
; LENGTH: 544  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-170-205E-15652

Query Match 95.0%; Score 2874.5; DB 27; Length 544;  
Best Local Similarity 98.7%; Pred. No. 1.2e-277;  
Matches 537; Conservative 0; Mismatches 0; Indels 7; Gaps 1;

Qy 26 MGATCAEPLQSVLWVKQORCAVSLPARALLRWRSFGPCGAGPGA-----DACSV 78  
Db 1 MGATCAEPLQSVLWVKQORCAVSLPARALLRWRSFGPCGAGPGAEGRGVFTDACSV 60

Qy 79 VSEIIAVEETDVHGKHOGSGKWQMEKPYAFTVHCVKRARRHRWKQAQVTFWCPPEQLCH 138  
Db 61 VSEIIAVEETDVHGKHOGSGKWQMEKPYAFTVHCVKRARRHRWKQAQVTFWCPPEQLCH 120

Qy 139 LWLQTLREMLEKLTSPKHLVFINPFGKGQGGKRIYERKVAFLFTLASITTDIIIVTEHA 198  
Db 121 LWLQTLREMLEKLTSPKHLVFINPFGKGQGGKRIYERKVAFLFTLASITTDIIIVTEHA 180

Qy 199 NOAKETLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLPSSLR 258  
Db 181 NOAKETLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLPSSLR 240

Qy 259 GIIPAGSDVCYSTVGTSDAETSAHLIVVGDLSAMDVSSVHNSLTLLRYSVSLGIGFY 318  
Db 241 GIIPAGSDVCYSTVGTSDAETSAHLIVVGDLSAMDVSSVHNSLTLLRYSVSLGIGFY 300

Qy 319 GDIIKDSKKRWGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFV 378  
Db 301 GDIIKDSKKRWGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFV 360

Qy 379 CRQSKQLEEEQKALYGLEAAEDVEEQQVVCVKFLAINATNMSCACRRSPRGLSPAHL 438  
Db 361 CRQSKQLEEEQKALYGLEAAEDVEEQQVVCVKFLAINATNMSCACRRSPRGLSPAHL 420

Qy 439 GDGSDLLIRKCSRNFRLIRHTNQDQDFTFVEVYRVKQFQTSKHMEDESDLK 498  
Db 421 GDGSDLLIRKCSRNFRLIRHTNQDQDFTFVEVYRVKQFQTSKHMEDESDLK 480

Qy 499 EGGKRFGHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIENPKP 558  
Db 481 EGGKRFGHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIENPKP 540

Qy 559 DSHS 562  
Db 541 DSHS 544

RESULT 14  
PCT-US02-24623-12  
; Sequence 12, Application PC/TUS0224623  
; GENERAL INFORMATION:  
; APPLICANT: EXELIXIS, INC.  
; TITLE OF INVENTION: SPKs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE  
; FILE REFERENCE: EX02-104  
; CURRENT APPLICATION NUMBER: PCT/US02/24623  
; CURRENT FILING DATE: 2002-08-02  
; PRIOR APPLICATION NUMBER: US 60/310,362  
; PRIOR FILING DATE: 2001-08-06

; PRIOR APPLICATION NUMBER: US 60/357,501  
; PRIOR FILING DATE: 2002-02-15  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 12  
; LENGTH: 481  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US02-24623-12

Query Match 85.6%; Score 2588; DB 1; Length 481;  
Best Local Similarity 100.0%; Pred. No. 4.8e-249;  
Matches 481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 82 IIAVEETDVHGKHOGSGKWQMEKPYAFTVHCVKRARRHRWKQAQVTFWCPPEQLCHLWL 141  
Db 1 IIAVEETDVHGKHOGSGKWQMEKPYAFTVHCVKRARRHRWKQAQVTFWCPPEQLCHLWL 60

Qy 142 QTLREMLEKLTSPKHLVFINPFGKGQGGKRIYERKVAFLFTLASITTDIIIVTEHANQA 201  
Db 61 QTLREMLEKLTSPKHLVFINPFGKGQGGKRIYERKVAFLFTLASITTDIIIVTEHANQA 120

Qy 202 KETLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLPSSLRIGII 261  
Db 121 KETLYEINIDKYDGI VCVGGDMFSEVLHGLIGRTORSAGVDQNHPRAVLPSSLRIGII 180

Qy 262 PAGSDVCYSTVGTSDAETSAHLIVVGDLSAMDVSSVHNSLTLLRYSVSLGIGFYGDI 321  
Db 181 PAGSDVCYSTVGTSDAETSAHLIVVGDLSAMDVSSVHNSLTLLRYSVSLGIGFYGDI 240

Qy 322 IKDSKKRWGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCRQ 381  
Db 241 IKDSKKRWGLARYDFSGKLTFLSHHCYEGTVSFLPAQHTVGSPRDRKPCRAGCFVCRQ 300

Qy 382 SKQLEEEQKALYGLEAAEDVEEQQVVCVKFLAINATNMSCACRRSPRGLSPAHLGDG 441  
Db 301 SKQLEEEQKALYGLEAAEDVEEQQVVCVKFLAINATNMSCACRRSPRGLSPAHLGDG 360

Qy 442 SSDLLIRKCSRNFRLIRHTNQDQDFTFVEVYRVKQFQTSKHMEDESDLKKEG 501  
Db 361 SSDLLIRKCSRNFRLIRHTNQDQDFTFVEVYRVKQFQTSKHMEDESDLKKEG 420

Qy 502 KKRFGHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIENPKPDH 561  
Db 421 KKRFGHICSSHPSCCCTVSNSSWNCDEVLHSPAIEVRVHCOLVRLFARGIENPKPDH 480

Qy 562 S 562  
Db 481 S 481

RESULT 15  
US-10-486-306-12  
; Sequence 12, Application US/10486306  
; GENERAL INFORMATION:  
; APPLICANT: EXELIXIS, INC.  
; TITLE OF INVENTION: SPKs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE  
; FILE REFERENCE: EX02-104C-US  
; CURRENT APPLICATION NUMBER: US/10/486,306  
; CURRENT FILING DATE: 2004-02-05  
; PRIOR APPLICATION NUMBER: US 60/310,362  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/357,501  
; PRIOR FILING DATE: 2002-02-15  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 12  
; LENGTH: 481  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-486-306-12

Query Match 85.6%; Score 2588; DB 30; Length 481;

Best Local Similarity 100.0%; Pred. No. 4.8e-249;			
Matches 481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	82	IIAVEETDVHGKHGSGKWQKMEKPYAFTVHCVKRARRHRKWAQVTFWCPEEQCHLWL	141
Db	1	IIAVEETDVHGKHGSGKWQKMEKPYAFTVHCVKRARRHRKWAQVTFWCPEEQCHLWL	60
Qy	142	QTLREMLEKUTSRPKHLLVFINPFGGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQA	201
Db	61	QTLREMLEKUTSRPKHLLVFINPFGGKGQKRIYERKVAFLFTLASITTDIIIVTEHANQA	120
Qy	202	KETLYEINIDKYDGIVCVGGDGMPSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGII	261
Db	121	KETLYEINIDKYDGIVCVGGDGMPSEVLHGLIGRTQRSAGVDQNHPRAVLVPSSLRIGII	180
Qy	262	PAGSTDCVCYSTVGTSDAETSALHI VVGDSLAMDVSSVHNHNTLLRYSVSLLYGFFYGDI	321
Db	181	PAGSTDCVCYSTVGTSDAETSALHI VVGDSLAMDVSSVHNHNTLLRYSVSLLYGFFYGDI	240
Qy	322	IKDSEKKRWJGLARYDFESGLKTFLLSHHCYSGTVSFLPAQHTVGSPRDKPCRAGCFVCRQ	381
Db	241	IKDSEKKRWJGLARYDFESGLKTFLLSHHCYSGTVSFLPAQHTVGSPRDKPCRAGCFVCRQ	300
Qy	382	SKQOLEEOKKALYGLEAAEDVEEWQVVCCKFLAINATNMNSCACRRSPRGLSPAHLGDG	441
Db	301	SKQOLEEOKKALYGLEAAEDVEEWQVVCCKFLAINATNMNSCACRRSPRGLSPAHLGDG	360
Qy	442	SSDLILIRKCSRNFLLRTHNTQQQDQDFTFVETVRVKKFOFTSKHMEDESDLKEGG	501
Db	361	SSDLILIRKCSRNFLLRTHNTQQQDQDFTFVETVRVKKFOFTSKHMEDESDLKEGG	420
Qy	502	KKFGHICSHSPSCCCTVSNSSNCDGEVLHSPAI EVRVHCOLVRLFARGIENPKPDSh	561
Db	421	KKFGHICSHSPSCCCTVSNSSNCDGEVLHSPAI EVRVHCOLVRLFARGIENPKPDSh	480
Qy	562	S 562	
Db	481	S 481	

Search completed: September 3, 2005, 04:25:25  
Job time : 349.06 secs

**This Page Blank (uspto)**